BEATRIZ TRINCHÃO ANDRADE DE CARVALHO

IMAGE-BASED APPEARANCE PRESERVATION

CURITIBA

2013

BEATRIZ TRINCHÃO ANDRADE DE CARVALHO

IMAGE-BASED APPEARANCE PRESERVATION

Tese de doutorado apresentada como requisito parcial à obtenção do grau de Doutora em Informática. Programa de Pós-Graduação em Informática, Setor de Ciências Exatas, Universidade Federal do Paraná.
Orientadores: Profa. Dra. Olga Regina Pereira Bellon e Prof. Dr. Luciano Silva

CURITIBA

2013

ACKNOWLEDGMENTS

First of all, I would like to thank my supervisors Prof. Dr. Olga Regina Pereira Bellon and Prof. Dr. Luciano Silva for all the support provided since I entered Imago Research Group. I learned a lot from you two, and I am truly grateful for all the excellent opportunities you made available for me during the years we worked together.

My special thanks go to Prof. Dr. Hendrik Lensch, for accepting me under his supervision during my internship in Tübingen University and for providing me a wonderful working experience there. Your interest and guidance were essential to the finalization of this thesis.

I would like to thank my colleagues from Imago Research Group and from Tübingen University's Computer Graphics Group for the support provided. I feel I am really lucky to have worked with such competent people and also to have made some good and lasting friends from both groups.

I am deeply grateful to Benjamin Resch, Tobias Häußler and Manuel Finckh, for the fruitful discussions and technical help during the last stage of this work. I am also grateful and highly indebted to Caroline Mazetto, for always being by my side whenever there was need, be it in scientific or in personal matters.

I would like to thank the committee members of my defense. You provided inestimable corrections and suggestions that improved this thesis. Financial support for this thesis was provided by CAPES-REUNI and by CAPES-PDSE (BEX 0157-12-0).

I thank my friends, who always made my life lighter. In particular, I am very grateful to my friends in Curitiba, for all the care they had with me and for the great moments we shared during the years I lived there.

I would like to thank my parents for their constant support and for always being close to me, regardless of geographical distances.

As it would not have been possible without his lovely company and never-failing encouragement, this thesis is dedicated to my husband André.

"Always look on the bright side of life!"

(Eric Idle)

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

$BRDF$ Bidirectional Reflectance Distribution Function

$BSSRDF$ Bidirectional Scattering-Surface Reflectance-Distribution Function

HDR High Dynamic Range

HVS Human Visual System

LDR Low Dynamic Range

MVS Multi View Stereo

$NRMSE$ Normalized Root Mean Squared Error

PDF Probability Density Function

$RMSE$ Root Mean Squared Error

SFM Structure From Motion

$SVBRDF$ Spatially-Varying Bidirectional Reflectance-Distribution Function

ULR Unstructured Lumigraph Rendering

ABSTRACT

The three-dimensional digital preservation of real objects comprises two main aspects: the preservation of the shape of the object and the preservation of its appearance. This thesis focuses on the image-based appearance preservation of real objects and provides a set of contributions on the theme. The first contribution consists in two groups of experiments, where each one of them targets one different image-based appearance preservation approach. These experiments are based in fundamental concepts related to the behavior of light and in a compilation of works that aim to preserve the appearance of real objects using different types of images. The first group of experiments attempts to disregard as much as possible the influence of the environment light. The second one goes one step further and considers a single light source. These experiments were the basis and motivation for the development of the main contribution of this thesis, which is a novel image-based appearance preservation method that considers the whole environment as a source of light. It presents as novelty the fact that it estimates the incoming light from the whole environment to each point in an object surface patch. At the best knowledge of this work, none of the current existing methods adopts this approach. Considering the whole environment as source of light allows flexible acquisition setups and, as it reproduces what happens in reality, potentially retrieves more reliable information about the incident lighting. This thesis presents this method and its application on real and synthetic environments. Conclusions about this work are presented and future research directions are discussed.

RESUMO

A preservação tridimensional digital de objetos reais compreende dois aspectos: a preservação da forma do objeto e a preservação de sua aparência. Esta tese tem como foco a preservação da aparência de objetos reais baseada em imagens e provê uma série de contribuições sobre o tema. A primeira contribuição consiste em dois grupos de experimentos, onde cada um trabalha uma abordagem diferente na preservação da aparência baseada em imagens. Esses experimentos são baseados em conceitos fundamentais relacionados ao comportamento da luz e em uma compilação de trabalhos que visam preservar a aparência de objetos reais usando diferentes tipos de imagens. O primeiro grupo de experimentos tenta desconsiderar ao máximo a influência da luz. O segundo vai um passo além e considera uma única fonte de luz. Estes experimentos são a base e motivação para o desenvolvimento da principal contribuição desta tese, que é um novo método de preservação da aparência baseado em imagens que considera todo o ambiente como fonte de luz. Ele apresenta como novidade o fato de estimar a luz vinda de todo o ambiente para cada ponto em uma região na superfície de um objeto. Até onde foi pesquisado neste trabalho, nenhum método existente adota essa abordagem. Considerar todo o ambiente como fonte de luz permite configurações flexíveis durante a aquisição e, já que reproduz o que acontece na realidade, recupera informações potencialmente mais confiáveis sobre a iluminação incidente. Esta tese apresenta este método e sua aplicação em ambientes reais e sintéticos. Conclusões sobre este trabalho são apresentadas e direções futuras de pesquisa são discutidas.

CHAPTER 1

INTRODUCTION

The three-dimensional (3D) digital preservation aims to generate a 3D representation of a real object for cultural or scientific purposes. It goes beyond the generation of a model because it comprises a complete process, which ranges from the data acquisition to the generation of the final model of objects, and is necessarily based on real data. During the preservation process, two key aspects must be considered: the preservation of the shape and the preservation of the appearance of the object (see Figure 1.1). The shape preservation involves capturing the geometry of the object and its representation. The preservation of the appearance involves capturing information about the materials that compose the object, in order to represent their reflectance properties.

The appearance preservation is a key factor in the reproduction of realistic 3D models, and goes beyond the representation of their surfaces colors. The color only represents the appearance of each point under specific lighting, and varies when lighting changes (see Figure 1.2). In its turn, the appearance preservation includes the reproduction of the behavior of a material under different lighting conditions. This is done by modeling the appearance of the 3D model surface as a function of its incident illumination through the use of reflectance models.

(a) (b)

Figure 1.1: 3D digital preservation of a ceramic artefact (available at Imago 3D Virtual Museum [32]): (a) preserved geometry; (b) preserved appearance.

(a) (b)

Figure 1.2: Pictures of a pine under different illumination settings.

Different reflectance models are used to reproduce the appearance of distinct materials. These models are usually generated from data representing the lighting and involved surfaces, captured using specific measurement devices or sensors. The complexity of the resulting models changes according to the properties of the materials to be reproduced. The simpler reflectance models generally cover at least four dimensions, representing the directions of incoming and outgoing light rays at a point. However, more parameters can be added to the models, depending on the properties the surface must present. For example, parameters can be added to account for spatially varying surfaces and temporal changes in the appearance of the material.

Sampling data to gather enough information to generate reflectance models is a cumbersome problem, as the material should be observed from many different positions and light settings to provide enough data to properly sample all the parameters defined by a reflectance model. Different approaches have been used to deal with this lack of information while retrieving realistic models. These approaches involve: to disregard the incoming light by reducing its influence; to consider only the light that has traveled directly from the light sources to the point being considered (direct lighting); or to assume that the light that comes from the whole environment influences object appearance.

1.1 Motivation

The appearance preservation of real surfaces is an important application of graphics processing concepts as it enables modeling the way a surface reflects light. The reflectance

models used to represent the appearance of an area on the surface of a real object should have similar behavior in the incidence of light as the same area in the original object. This makes it possible to render this area under novel camera positions or to render arbitrary shapes using the materials preserved by the reflectance models. This knowledge can also be used to gather information about materials from which the object is composed. It is noteworthy that the appearance preservation can be directly used in the digital preservation of heritage, the automotive industry, game development, movie production, among others.

The first two approaches described previously in this chapter present plausible results at the cost of drawbacks caused by the loss of the environment lighting contribution and limitations in the environment appearance. Not considering the incoming light is a straightforward approach, that enables fast results with reduced efforts. However, this approach is error-prone, as strong light sources, when projected on the object surface, influence the resulting color and reduce the precision of the generated models.

Considering direct lighting provides more realistic results than the first approach, as some lighting information is taken in consideration during the generation of the reflectance models. As the direct lighting must be the only source of light in this approach, it presents limitations in the environment setup like the requirement of dark environments or light position restrictions during data acquisition. Furthermore, it is not possible to eliminate the influence of the environment completely as the scene surfaces always reflect some light from the light sources and influence the observed colors.

To consider the whole environment as source of light reproduces what happens in reality, having the potential to retrieve the most precise results. This approach provides more flexible acquisition setups, as there are no restrictions about the directions the light comes from. This is a more complex approach because besides the fact that it is necessary to capture and represent the light that comes from the whole environment, there is far more information to deal with during the generation of reflectance models. For these reasons, the third approach is not as investigated as the first two ones; at the

best knowledge of this work, few works estimate or capture the lighting that comes from the whole environment and use it to calculate reflectance models.

However, none of these approaches captures the light from the whole environment for each point in the object surface. The solution proposed by Haber [27] estimates the incoming light using only images of the object, while Häußler [31] estimates the incoming light using an image of the environment captured using a fisheye lens. This image is used to represent the incoming light from the environment to a surface patch, and only one image of the environment is captured and used to estimate the light from the environment to every point in the patch.

However, in the real world the incoming light changes from point to point, sometimes drastically. For instance, consider the two sides of a line defined by a shadow on a surface patch. The incoming light at the shadow side is much weaker than the incoming light in the other side. This is an important characteristic that greatly increases the complexity of the problem, as an environment light map should be captured for each point in the patch.

Considering this scenario, the motivation of this work is to investigate appearance preservation approaches and show that by considering the influence of the incoming light from the whole environment one could achieve more realistic results and a more flexible acquisition setup. As the observed color of a surface point is the result of the light that comes from the whole environment and is reflected to the direction of the observer, this approach reproduces this phenomena with more fidelity than the other ones. Also, considering the whole environment means less or no restrictions regarding the environment appearance, the camera positioning and the characteristics and number of light sources. As discussed in this section, those features are restricted in the other two approaches.

To this end, the hypothesis of this thesis is that it is possible to develop a method that estimates the incoming light from the whole environment to *each point* in an object surface patch, generating space variant reflectance models for this patch that can simulate the surface appearance. It is assumed that the lighting does not change strongly during time.

This way, the following goals were defined to achieve this purpose:

- To investigate methods that aim to preserve the appearance of objects.
- To acquire experience in appearance preservation and apply the investigated techniques by developing methods that address the first two preservation approaches.
- To propose a new method to preserve the appearance of real objects considering the whole environment as source of light, following the hypothesis delineated in this section.
- Conduct tests to show the validity of all the developed methods.

1.2 Methodology and Contributions

Given an area on the surface of an object inside a scene, this thesis focuses on the preservation of the appearance of this area. This is done by estimating its reflectance properties, that describe the behavior of the light that comes from the environment and is reflected by the object's surface. The data used to generate these models can be acquired by specialized equipment such as a gonioreflectometer. However, as will be discussed in Subsection 2.3.3, this is usually a time and resource consuming approach.

This thesis reviews and focuses on image-based appearance preservation methods. In the last decades there has been an increasing investment in the use of imaging sensors to model the reflectance of objects. This approach stands out as a straightforward, easy and cost-effective way to capture data that can be used to model the reflectance of surfaces. In addition, they provide a portable data acquisition step that can be used to capture different kinds of surfaces.

Considering the three appearance preservation approaches outlined in this chapter and the defined goals of this thesis, the following methodology has been adopted:

1. Study of concepts about the appearance of materials.
2. Review of image-based works in the literature that aim to capture and represent the appearance of an object or material.

3. A set of experiments on the first and second approaches.

4. Development of a new image-based appearance preservation method using the third approach.

5. Application of the new method in real and synthetic settings.

A set of contributions can be derived from this thesis. First, a literature review and discussion about the state-of-the-art appearance preservation approaches are presented. Also, two sets of experiments were performed using the first two approaches described in this section. These experiments consist in developing methods to preserve the appearance of surfaces using each of these approaches. They are performed to apply the reviewed methods in a practical and known context, and to identify good project choices in appearance preservation. The methods developed during the experiments were evaluated and applied to real objects.

Based on the knowledge acquired from the previous contributions, a main contribution is presented, that is a novel image-based appearance preservation method that considers the incoming light from the whole environment. This method was developed to prove the hypothesis delineated in Section 1.1, and shows how the third approach can be adopted to provide a more flexible acquisition setup and consistent results while considering the light that comes from the whole environment to each point.

1.3 Chapter Overview

This thesis is organized as described below:

Chapter 2 - Background: this chapter presents an overview of the three image preservation approaches and fundamental concepts related to the work developed in this thesis. Basic elements on the comprehension of the behavior of light are described, as well as ways to use these elements to model the appearance of surfaces. It also describes a technique that enables the storage of greater sampling between the darkest and brightest areas in an image.

Chapter 3 - Image-Based Appearance Acquisition and Representation: this chapter makes a review of works in the literature that aim to capture and represent the appearance of real objects. This review focuses on image-based methods, so different types of images are considered. This chapter also presents a discussion containing a comparison between the reviewed methods and considerations about them.

Chapter 4 - Experiments: two main experiments were performed in order to understand and apply appearance preservation concepts. Both are appearance preservation methods, where the first one reduces at most the influence of lighting and the second one assumes direct lighting, using a single light source. Both of them are described, applied and evaluated in this chapter.

Chapter 5 - An Image-Based Preservation Method: During the experiments made in this thesis, it was noticed that a method that could consider the influence of the incoming light from the whole environment could obtain better results and provide a more flexible acquisition setup. This chapter presents a new method using this configuration, as well as the results obtained during its application on synthetic and real settings.

Chapter 6 - Conclusions: this chapter provides final remarks about this thesis and delineates future research directions.

CHAPTER 2

BACKGROUND

The appearance of an object is the result of a combination of different factors which include lighting, material characteristics, the position of the viewer and the environment where the object is located. When acquiring and reproducing the appearance of objects in virtual environments, one common approach is to reduce the complexity of the problem through assertions about these factors. This makes it possible to capture and digitally reproduce the appearance of an object using feasible requirements.

Therefore, it is very important to understand this phenomenon in order to recognize all the factors involved and use them efficiently. To this end, this chapter presents a brief description of the problem addressed in this thesis (Section 2.1), and a compilation of important notions regarding the characterization of light flow through space (Section 2.2) and how they can be used to represent the appearance of materials under different angles of incident light (Seção 2.3). Section 2.4 describes High Dynamic Range (HDR) Images, that enable the generation of images with a greater sampling between their darkest and brightest areas.

2.1 Problem Overview

As described in Chapter 1, one way to categorize appearance preservation methods is on how each one considers the incoming light from the environment, adopting one of three approaches. In the approach that disregards the incoming light, works like the ones developed by Andrade *et al.* [3], Farouk *et al.* [23] and Rushmeier and Bernardini [59] aim to reduce at most the influence of lighting in order to get a color that is closest to an average as possible. Generic reflectance models are then manually set using these average colors and taking into consideration the appearance of the original object.

When only direct lighting is considered, the challenge is to find the main light sources in the environment and to estimate their influence over the color observed in the images taken by a camera. This way, through the processing of these images and some assumptions about the surface properties, it is possible to estimate the incident lighting from the environment and model the way the surface reflects it. However, the incident light in a surface point is not integrated only over the rays that come from the light sources. Actually, it is integrated over all the rays that come from the environment and reach that point. In their turn, these rays are also generated in the same fashion, integrating light that comes from the whole environment to its starting point and reflecting it in the ray's direction (considering it is not a light source). This is a difficult problem, as the tracing of these rays back to the light sources generally makes the problem impossible to be solved analytically (Pharr and Humphreys [57]).

This way, methods that assume direct lighting must neutralize the influence of the environment. Levoy *at al.* [45] take images with the direct lighting turned on and off and assume that the difference between these images represents an image only illuminated by the direct lighting. Alldrin *et al.* [1] uses a dark room and reduce ambient illumination by acquiring ambient images, retrieved by blocking the light source so that the test object is in shadow. Works like the ones proposed by Lensch *et al.* [43], Marschner *et al.* [50] and Sato *et al.* [61] cover the environment with black fabric or paint it in black in order to minimize this influence.

To embrace the whole environment, using all the light that comes from it, is the most challenging approach. Despite being the one that potentially models the light reflection with more precision, this is a less explored approach as its implementation is more complex. Haber *et al.* [27] follow this approach by estimating the incoming light using images of the object in different light settings. Häußler [31] proposes an approach based on the utilization of one environment light map captured using fish eye lens.

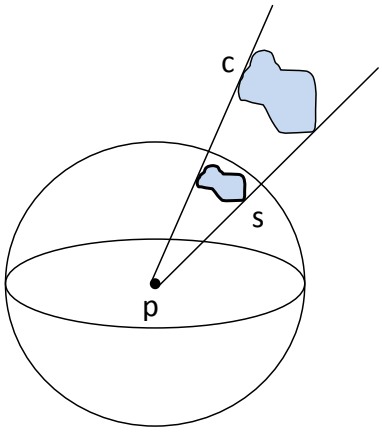


Figure 2.1: Object c and its solid angle s as seen from a point p.

2.2 Radiometry

Radiometry concerns the study of the propagation and reflection of electromagnetic radiation in an environment. According to Pharr and Humphreys [57], the light visible to humans refers to wavelengths of electromagnetic radiation from approximately 370nm (bluish colors) to 730nm (reddish colors). This section introduces some quantities that are used in the representation of material appearance, based on the descriptions made by Weyrich *et al.* [76], Pharr and Humphreys [57] and Hecht [29].

The first measurement is the amount of light flowing through some area per unit time. This value is know as *radiant flux* or *power* (Φ), and is measured in watts (W) or joules per second (J/s). This value represents a light which is emitted uniformly in all directions.

When light flows in a non-uniform way, it is necessary to quantify it in relation to a direction. This is represented through the radiant flux Φ emitted per solid angle. *Solid angles* are a way to measure the area occupied by a tridimensional region as viewed from a given point p. This way, given some 3D region in space, its solid angle is the area of its projection over a unit sphere around p (see Figure 2.1). The solid angles are measured in *steradians* (sr), and are usually represented by the symbol ω. Now it is possible to define the *intensity or directional radiant flux* (I):

$$I = \frac{\mathrm{d}\Phi}{\mathrm{d}\omega}. \qquad [W \cdot sr^{-1}] \tag{2.1}$$

It is important to stress that many radiometric definitions are written in differential form in order to emphasize that the involved values are infinitesimally small. As an example, the directional radiant flux is the limit of the power per solid angle ω, when the solid angle area shrinks to zero.

Irradiance (E) is the amount of light illuminating a surface, and is defined as the average energy per unit area per unit time, that is equal to the radiant flux per unit area A:

$$E = \frac{\mathrm{d}\Phi}{\mathrm{d}A}. \qquad [W \cdot m^{-2}] \tag{2.2}$$

Based on this definition, two laws emerge:

- Inverse square law: the irradiance from a point source is proportional to the inverse square of the distance;
- Cosine law: tilting the surface in relation to the light source changes the irradiance in proportion to the cosine of the angle between the surface normal and the direction of the light source.

Radiance (L) is the flux density per unit area A, per solid angle ω. However, instead of directly using the area A, the convention used in this definition is to use the projection of the area in a plane that is perpendicular to the viewer direction (*projected area* $A^{\perp}$):

$$L = \frac{\mathrm{d}\Phi}{\mathrm{d}\omega \ \mathrm{dA}^{\perp}}. \qquad [W \cdot m^{-2} \cdot sr^{-1}] \tag{2.3}$$

According to Weyrich *et al.* [76], the irradiance on a sensor is proportional to the radiance of the imaged surfaces, where the optical system integrates the radiance of the area defined by the solid angle subtended by the device aperture. Therefore, the irradiance in each pixel is converted to electrical signals and digitized, such that pixel values in a digital image are proportional to the incoming radiance.

2.2.1 Incident and Exitant Radiance Functions

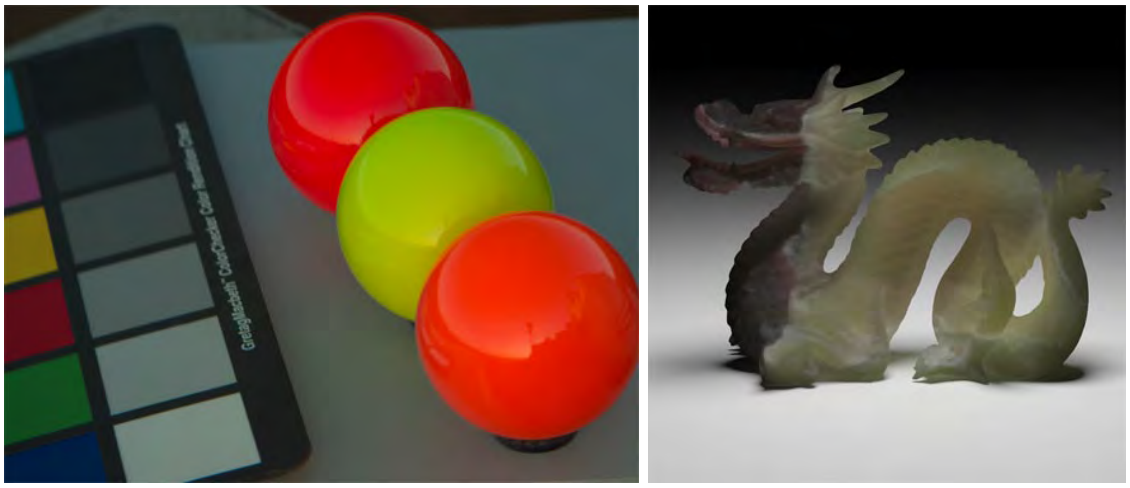
When using radiance values in equations, it is useful to make a distinction between the radiance that arrives in a point and the radiance that leaves this point. Pharr and Humphreys [57] use the notion of incident and exitant radiance functions to distinguish between these two cases. This way, given a point p, the radiance arriving at that point from the direction defined by ω defines the incident radiance function $L_i(\mathrm{p}, \omega)$. Similarly, the function that describes the radiance leaving p in the direction defined by ω is denoted as the exitant radiant function $L_o(\mathrm{p}, \omega)$. In both cases the vector defined by ω points away from p. However, it is important to add that some authors use a notation where this vector points to p in L_i.

2.3 Bidirectional Reflectance Distribution Functions (BRDFs)

The first step in the reproduction of what is seen in the real world is the analysis of how different materials interact with light. This interaction can be given in different ways: the material can reflect light, absorb it, or can display more complex phenomena such as fluorescence or subsurface scattering (Figure 2.2). Physically-based models are used to reproduce the appearance of materials under lighting. For different materials, different models are used and the complexity of these models varies with the properties of the material being reproduced.

Bidirectional Reflectance Distribution Functions (BRDFs) describe the reflectance of a point p on a surface through the quotient of the exitant radiance L_o and the incoming irradiance E_i at this point. The radiance L_o is emitted in a direction defined by ω_o and the irradiance comes from a direction defined by ω_i (see Figure 2.3), which usually makes BRDFs functions of four variables when they are constant over the surface points:

$$f_r(\omega_i, \omega_o) = f_r(\theta_i, \phi_i, \theta_o, \phi_o) = \frac{\mathrm{d}L_o(\omega_o)}{\mathrm{d}E(\omega_i)}. \tag{2.4}$$



(a) (b)

Figure 2.2: Examples of phenomena observed in the behavior of different materials under illumination: (a) fluorescence (Hullin *et al.* [30]); (b) subsurface scattering (Weyrich *et al.* [76]).

The parameters θ_i, ϕ_i and θ_o, ϕ_o are respectively the polar coordinates of ω_i and ω_o. When the BRDF changes over the surface, a parameter containing the point position on the surface must be defined ($\mathrm{p}(x, y)$), adding two more parameters to the equation:

$$f_r(\mathrm{p}, \omega_i, \omega_o) = \frac{\mathrm{d}L_o(\mathrm{p}, \omega_o)}{\mathrm{d}E(\mathrm{p}, \omega_i)}. \tag{2.5}$$

Based on the definitions of radiance and irradiance presented in Section 2.2, the incoming irradiance at a point p from a direction $E(\mathrm{p}, \omega_i)$ can be defined as a function of the incident radiance $L_i(\mathrm{p}, \omega_i)$ at that point:

$$\mathrm{d}E(\mathrm{p}, \omega_i) = L_i(\mathrm{p}, \omega_i) cos\theta_i \mathrm{d}\omega_i. \tag{2.6}$$

Using Equation 2.6 in the BRDF definition of Equation 2.4, it is possible to define the BRDF as a function of the incident and exitant radiances.

$$f_r(\mathrm{p}, \omega_i, \omega_o) = \frac{\mathrm{d}L_o(\mathrm{p}, \omega_o)}{L_i(\mathrm{p}, \omega_i) cos\theta_i \mathrm{d}\omega_i}. \tag{2.7}$$

This definition establishes grounds for the *reflection equation* (Pharr and Humphreys [57]), shown in Eq. 2.8. This equation can be obtained by isolating $\mathrm{d}L_o(\mathrm{p}, \omega_o)$ and integrating

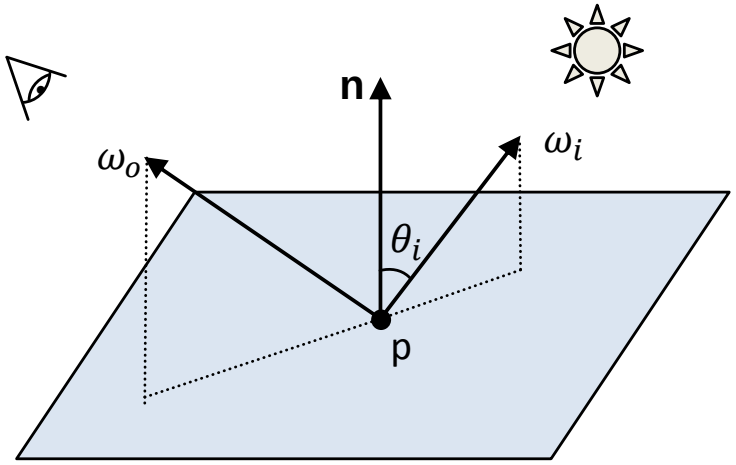


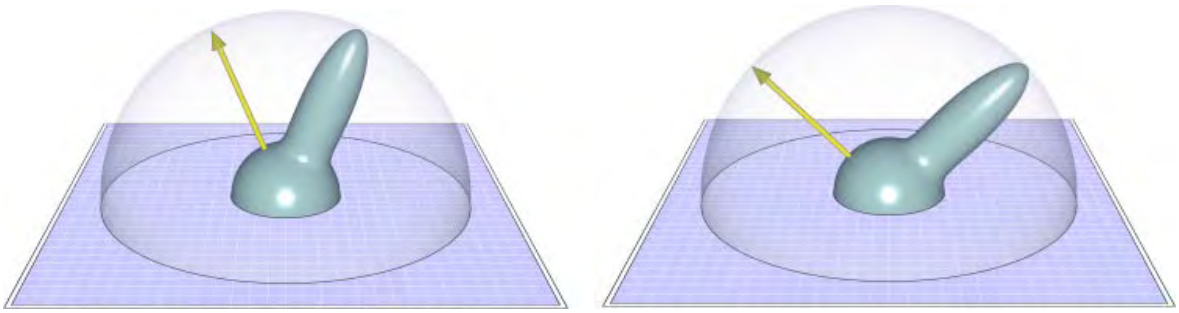
Figure 2.3: The BRDF is a function of the incoming radiance from ω_i and the exitant radiance in a direction defined by ω_o. The angle θ_i is defined by the normal vector $\mathbf{n}$ and the direction defined by ω_i.

the equation over the hemisphere $\mathcal{H}^2(\mathbf{n})$ of incident directions around the normal vector $\mathbf{n}$ of the point p. This equation uses the absolute value of the $cos\theta_i$ term to account for normals that are not reoriented to lie on the same side of the surface as ω_i, and to assume that the surface normal always points outside the surface.

$$L_o(\mathrm{p}, \omega_o) = \int_{\mathcal{H}^2(\mathbf{n})} f_r(\mathrm{p}, \omega_i, \omega_o) L_i(\mathrm{p}, \omega_i) |cos\theta_i| \mathrm{d}\omega_i. \tag{2.8}$$

The distribution of transmitted light can be similarly described by using the surface's *Bidirectional Transmittance Distribution Function* (BTDF). The BTDFs are denoted by $f_t(\mathrm{p}, \omega_i, \omega_o)$, where ω_i and ω_o are in opposite hemispheres around p. When used together with the BRDFs, they define the *Bidirectional Scattering Distribution Functions* (BSDFs), expressed by $f(\mathrm{p}, \omega_i, \omega_o)$. Using the BSDFs, the *scattering equation* (Pharr and Humphreys [57]) can be defined using the whole sphere S^2 of directions as domain:

$$L_o(\mathrm{p}, \omega_o) = \int_{S^2} f(\mathrm{p}, \omega_i, \omega_o) L_i(\mathrm{p}, \omega_i) |cos\theta_i| \mathrm{d}\omega_i. \tag{2.9}$$



(a) (b)

Figure 2.4: Goniometric plots of a BRDF slice corresponding to an incident light direction, represented as the yellow vector (Weyrich *et al.* [76]).

2.3.1 BRDF Visualization

BRDFs can be visualized through goniometric plots (see Figure 2.4). In this plots BRDFs are illustrated as a function of one incident light direction and of all possible directions of reflection. The reflected radiances in each direction are represented as points on the surface of the plot. It can be seen from the BRDF in Figure 2.4 that part of the incident light is reflected equally in all directions (part of the sphere with constant radius). However, there is a bump on the surface, indicating a direction in which the radiance values are higher.

If the direction of incidence of light in Figure 2.4(a) is changed, one can see that the direction of the bump changes, while the constant portion of the function does not change (Figure 2.4(b)). This happens because in this function radiance is higher toward the ideal reflection of the incident light. This feature characterizes the *specular peaks*, usually observed in objects with shiny surfaces. In turn, the portion of the sphere with constant radius is the *diffuse reflection* related to the color of the material.

There are generalizations of BRDFs that account for the complex materials mentioned in the beginning of this section. As examples, Weyrich *et al.* [76] describe the SVBRDFs (Spatially-Varying Bidirectional Reflectance Distribution Functions) and the BSSRDFs (Bidirectional Scattering Surface-Reflectance-Distribution Functions). The SVBRDFs represent spatially varying BRDFs on the surface of the object. A SVBRDF is usually expressed as a function of six variables (four to represent the BRDF for each point and two to represent the point position on the object surface). The BSSRDFs consider

the diffusion of light under the surface, and can be expressed using eight variables (six variables to represent the SVBRDFs plus two to represent the point on the surface where the incident light leaves the object).

2.3.2 Properties and Characteristics of BRDFs

According to Weyrich *et al.* [76] and Pharr and Humphreys [57], to be considered physically plausible, a BRDF must present two properties: *energy conservation* and *Helmholtz reciprocity*. Energy conservation is guaranteed when a surface never reflects more light than the total that was focused on it. Thus, all light must be reflected or absorbed, and no light should be created during the reflection. The Helmholtz reciprocity is observed when the value of the BRDF does not change when the incident and exitant light directions are switched.

Depending upon the materials that are represented, the BRDFs may present some characteristics:

- Isotropy: the BRDF does not change if the incident and exitant light directions are rotated at the same rate around the surface normal. It is important to note that this feature allows the simplification of a BRDF to three dimensions, as it comes to depend only on the difference between the azimuthal angles of incidence and exitance.

- Anisotropy: the inverse of isotropy, usually observed in materials like brushed steel, hair and velvet.

- Retro-reflection: the light is reflected more strongly in the direction of the incident light. This is the case of materials used in the painting of traffic signs.

2.3.3 BRDF Modeling

The modeling of real BRDFs can be split into two main steps: sample acquisition and BRDF representation from these samples. There are different processes of acquisition

and representation of BRDFs. Regarding the acquisition, specific equipment can be used, such as gonioreflectometers, or different types of imaging sensors. The gonioreflectometer is a device that measures the BRDF of a point on a surface. It is basically composed of a photo-detector and a light source, where the photo-detector measures the light reflected from the light source on a point. The measurement of the light reflected from a set of different photo-detector and light source positions samples the BRDF of this point. The gonioreflectometer is generally used to describe material samples, since the measurement over all points on a surface is laborious and time consuming.

Due to the gonioreflectometers practical limitations, the use of different types of images is being considered to generate BRDFs. As examples, the utilization of color images, HDR images (see Section 2.4), polarized images, depth images and reflectance images have been researched. Frequently, more than one of these types of images are used together. Chapter 3 presents recent works that make use of different types of images with the goal of capturing BRDFs.

From the collected samples, the next step is to generate models that simulate the appearance of the material under arbitrary illumination. For this task it is necessary to define how the BRDF is represented. One way to represent a BRDF is through tabular data, where a large number of samples is captured and stored for direct access using the values of the incidence and reflection angles. However, the capture of samples for all the possible incidence and reflection angles is unfeasible, since there are practical constraints such as the positioning of the device in all positions of incidence and reflection of light, the amount of time required to do so, and the great quantity of data generated. Thus, what normally happens is the reduction or interpolation of these data to obtain computationally feasible representations.

Another approach is the formulation of analytical BRDF models, detailed in Subsection 2.3.3.1. Other approaches include obtaining parameters for BRDF models by representing the reflectance functions as spherical harmonics or by their decomposition and description in the frequency domain. Some studies also assume functions where no

model is assumed and the reflectance at each point is characterized as a linear combination of base functions.

From a set of observations of the same material, the difference between each sample and the value estimated by the model for this sample can be minimized in order to obtain the model parameters that ideally represent the material characteristics. It is also possible to define a material as a linear combination of known BRDF models, which would work as a set of basis BRDFs. In this case, the difference between each sample and the linear combination of values estimated by the basis BRDFs would be used to retrieve the coefficients that minimize the difference between the sample and the linear combination. Chapter 3 presents several works that use the techniques of representation mentioned in this section, and a discussion about these approaches is presented in Section 3.5.

2.3.3.1 Analytical BRDF Models

Analytical BRDF models represent the appearance of materials under incident light, and beyond the values of the angles, they depend on other variables, according to their formulation. These variables define certain material characteristics, such as roughness or proportions of diffuse and specular reflection. According to Ngan [56], analytical BRDF models can be roughly divided into two classes: empirical models and physically-based models.

According to Lensch [41], empirical models are typically not physically motivated, but provide a class of functions that can be used to approximate BRDFs. Examples of this class include the Blinn-Phong model [10], the Lafortune model [37] and the Ward model [74].

Physically-based models are derived from principles of physics. The Torrance-Sparrow model [67], the He model [28], the Cook-Torrance model [14] and the Ashikhmin-Shirley model [7] are examples of this class. These works are based on the assumption that a surface is composed of microfacets whose properties are defined by distribution functions.

Weyrich *et al.* [76] and Ngan [56] present a detailed overview of the most popular analytic BRDF models. For more details regarding each model, the reader is referred to

the original papers. An evaluation and comparison of several analytical models in terms of their ability to reproduce measured BRDFs was performed by Ngan *et al.* [54] (see Section 3.5).

2.4 High Dynamic Range Imaging

All the radiometric measurements have corresponding photometric measurements. The *luminance* of a surface measures how bright the distribution of visible light at each wavelength (*spectral power distribution*) appears to a human observer, and is related to the spectral radiance of the surface (Pharr and Humphreys [57]).

A real scene can contain a high range of contrast and luminance. Due to its receptors, the human visual system (HVS) has the ability to adapt to different scenarios. Thus, the human eye can distinguish at a single frequency about 10000 different colors. With some time to adapt, it can reach values of 14 orders of magnitude (Gonçalvez *et al.* [25]). This relationship between the maximum and minimum values defines the *dynamic range*. The dynamic range captured by the human visual system is called *High Dynamic Range* (HDR).

Most imaging devices are still quite limited regarding the dynamic range, capturing Low Dynamic Range (LDR) images. Therefore, when the specular reflection in a region is very strong relative to the dynamic range of a camera, there is pixel saturation in this region. When they exceeded the maximum dynamic range of the camera, they are registered with the maximum intensity value. All colors and details contained in this region are recorded as white, with a loss of possibly relevant information. Likewise, the darker regions of a brightly lit scene may be located below the dynamic range of the image if details about the brighter region should be imaged. This way, the darker regions are recorded with the minimum intensity values (black), despite of their real color. In both cases, in order to map all the colors inside a low dynamic range, a color is recorded with intensity values that are relative to the other colors in the scene.

Figure 2.5: HDR imaging. First line: images taken with different exposure times; second line: resulting tone mapped HDR image (Weyrich *et al.* [76]).

Thus, new alternatives to capture, store and view images have been researched in order to approximate imaging sensors to the human viewing experience. In this context arise HDR images, which aim to reproduce the high dynamic range visible by the HVS.

To get the HDR image of a scene it is necessary to obtain the maximum color variation presented there. To this end, normally several photographs are captured with different exposure times. Thus, regions with specular peaks have their color captured with better quality at lower exposure times, while the dark regions of the photo require larger exposure times to compensate for the low luminance. Thus, different dynamic ranges are sampled in each photo, and the combination of these photographs will result in an HDR image (see Figure 2.5).

To store this new range of values without loss of information, new formats were developed. According to Gonçalves *et al.* [25], some of the most used are: RGBE [73], LogLuv TIFF [39], and OpenEXR [11]. To display HDR images on regular LDR monitors, a *tone mapping* process is performed. This process maps the HDR values to values inside the lower dynamic range of the monitor.

2.5 Remarks

This chapter presented radiometric concepts related to the comprehension of light behavior and ways to use these elements to model the appearance of surfaces through bidirectional reflectance distribution functions. It also introduced the high dynamic range imaging, a method that enables the storage of a greater range of luminance and contrast.

The concepts presented in this chapter will be used throughout this thesis. Works on image-based BRDF generation are reviewed in Chapter 3, including HDR image-based methods. They are also fundamental in the understanding of the decisions made during the development of the methods presented in Chapters 4 and 5.

CHAPTER 3

IMAGE-BASED APPEARANCE ACQUISITION AND MODELING

The acquisition and modeling of the appearance are two essential stages in the digital preservation of real objects. In the *acquisition stage* data is captured from the object (and usually from its environment) through the use of sensors. The data obtained is then used in the *modeling stage*, where techniques are used to calculate BRDFs or reflectance models and store them efficiently.

The purpose of the acquisition stage is to estimate the radiance emitted by the materials contained in the object. This value is typically estimated using a gonioreflectometer, equipment described in Subsection 2.3.3. Despite being a device that provides consistent measures of BRDFs at one point, the gonioreflectometer has low feasibility in measuring BRDFs of full objects. Thus, in the last decades there has been a considerable investment in the use of imaging sensors as capture device in the acquisition stage, and in the development of appearance modeling methods that are compatible with these devices.

This chapter presents a compilation of works that aim to capture and model the appearance of a material with the use of images. Several characteristics can be chosen to categorize such methods. Some examples are the kind of imaging they use, the type of reflectance they aim to model or how the reflectance is represented. In this chapter, the works are clustered according to the type of images used as input in the acquisition stage:

- Color images (Section 3.1);
- HDR images (Section 3.2);
- Color or HDR images (Section 3.3);
- Polarised images (Section 3.4);

- Spectral images (Section 3.4).

This choice categorizes well the various image-based methods for two main reasons: the type of the input image is closely related to the reflectance model chosen; and it provides a clear overview of the imaging possibilities and how they influence the modeling stage. Nevertheless, a comparison between these methods regarding further key aspects is discussed in Section 3.5 and conclusions are presented in Section 3.6. It is important to notice that depth images are not listed as a category above. This was intentionally done because, as will be detailed in this chapter, most works on image-based appearance preservation use depth information at some level.

3.1 Color Images

This section compiles relevant works that use color images as source of information to model the appearance of an object. These works use color images taken by photographic cameras, and geometric information is obtained through the processing of these images or from the use of depth acquisition devices. As regular color images have low dynamic range, photometric calibration methods are often required to find reliable reflectance or luminance values.

Levoy *et al.* [45] describe the digital preservation of works made by the artist Michelangelo. The capture of the appearance is made using color images captured by a digital camera. The geometry is measured using a scanner, and the obtained normal vectors are used as parameters for converting the color captured by the camera into reflectance values for each model vertex.

This work uses a highly controlled environment where the position and properties of the illumination (a halogen lamp treated as point light source) are known. The lamp is used to project uniform light over the object surface. During the image capture, two images are obtained from each perspective: one with the light on and another with the light off. The authors assume that the subtraction of both images results in an image

only illuminated by the lamp. The camera geometric distortion and the radiometric effects (*e.g.* angular and spatial non-uniformity) are adjusted.

The illumination and geometry are known, and the irradiance at the point and the reflected radiance towards the camera are estimated and used to compose a surface reflectance function. As the radiant intensity of the light source and the scale of the radiance estimated by the function are unknown, a second reflectance function based on the images of a white card is defined. This function is divided by the reflectance function of the object. This way, the parameters in common are eliminated, and by using the white card standard reflectance values it is possible to estimate the absolute reflectance of the surface.

This way, each vertex of the model has a set of reflectance values obtained through different photographs. A set of characteristics is used to define how each point will contribute in the final texture. As an example, to decrease the sensor saturation, observations near the reflection of light are discarded. Thus, the final reflectance in a point is obtained through a weighted sum of the reflectance values estimated for that point using different images.

This work does not consider inter-reflections or specular reflections, and a Lambertian surface is assumed. As the objects preserved in this work are mostly composed of antique marble, these considerations generate plausible results. However, the discard of information near the reflection direction is a problem because important data about the surface reflection is lost.

The authors observe that many details that are too small to be sampled by the scanner are not captured and therefore will not be considered in the irradiance calculation, which makes them appear as changes in diffuse reflection. The authors also note that despite having good visual results (see Figure 3.1), one should be cautious when using the results to scientific analysis.

Rushmeier and Bernardini [59] developed an approach which integrates reflectance values obtained in different images using a weighted sum. Here, the geometry of the object is captured at low resolution, using a system based on cameras and lighting. Five

(a) (b)

Figure 3.1: David (Levoy *et al.* [45]): (a) picture; (b) image of the rendered model.

lamps are attached to this system, and for each view of the object five images are taken with each lamp turned on at a time. This set of images is used to calculate more accurate normal vectors for the model by using photometric stereo.

It is assumed that the object has a Lambertian surface, and can have its appearance modeled by a BRDF. Thus, the new normal vectors are used to calculate the reflectance of the surface along with values of the position and direction of light. The authors do not describe how these last two values are obtained, but as the illumination position is fixed relative to the camera, it is understood that they are manually estimated.

For each set of five images with different illumination, five different reflectance values are obtained for each point. These values are adjusted in relation to each other in order to compensate for variations in the intensity of the light source. The adjusted images are combined to obtain a corrected RGB image for each camera position. As the reflectance is relative to other pixels in the same image, this approach corresponds to a relative value. To obtain the absolute value of the reflectance, measurements are made at some points using a device called Colortron. These measurements are used to adjust the relative reflectance of the images so that they represent visible values in the desired spectrum. This way, for each point of the object there is a set of observed reflectance values.

These values must be merged in order to define the final color. To avoid errors caused by small misalignment of images, the reflectance observed for each point is the average of the reflectances observed in its neighborhood. In order to minimize the difference

(a) (b)

Figure 3.2: Pietà (Rushmeier and Bernardini [59]): (a) overlapping photographs; (b) image of the rendered model.

between the reflectances detected in all corresponding observations, a global adjustment of the reflectance values is performed. After the adjustment of all images, correction values are globally applied to the red, green and blue, using measurements from Colortron. Figure 3.2 presents the method results on Michelangelo's Pietà.

This method was later used by Farouk *et al.* [23] to preserve assets of the Egyptian Museum in Cairo. As Levoy *et al.* [45], the method calculates average reflectances which minimizes specular reflection. In a case described by Farouk *et al.*, the specular reflection was manually set to simulate the gold plating of a piece.

While Levoy *et al.* [45] and Rushmeier and Bernardini [59] only preserve the diffuse reflection of objects, Sato *et al.* [61] present an approach to model the specular and diffuse reflection of objects using a set of color and depth images. The method creates complete models with non-uniform reflectance properties. The geometry reconstruction is described in the work, and in terms of reflectance modeling, the basic idea of the article is to use the Torrance-Sparrow model [67] to separate specular and diffuse components from a sequence of observations of a point on the object, represented as an array of RGB colors.

This array is presented as part of the Torrance-Sparrow model. Some of the model values are assumed or calculated, so that it is possible to solve the system and divide the

(a) (b)

Figure 3.3: Mug (Sato *et al.* [61]): (a) photograph, (b) image of the rendered model.

model into two parts: one representing the specular reflection, and another the diffuse reflection. From this separation, the authors calculate the diffuse component of the model.

As the specular component is observed only from restricted angles, it cannot be calculated in some regions of the object. Thus, the authors also propose a method to interpolate the specular component on the surface of the object. A set of metrics is defined to choose the vertices of the object with the greatest specular influence. The specular reflection parameters in these vertices are interpolated on the faces of the object.

The method appears to be stable and shows consistent results (see Figure 3.3). To improve the specular detection, acquisition planning methods could be used in this context in order to ensure maximum specular reflection capture.

Marschner *et al.* [50] present a process to measure the BRDF of homogeneous surfaces. The method requires cylindrical or spherical objects, although curved or almost convex objects can be used if their geometry is known. The idea of the method is to fix a point light source and capture multiple pictures of the object at angles covering an entire hemisphere of observations (see Figure 3.4).

The camera must be submitted to radiometric and geometric calibrations. The geometric calibration calculates the relative positions of the light source, the object and the camera for each image. This calibration is done through the use of patterns placed in the scene. The radiometric calibration determines the irradiance due to the light source and the relationship between pixel values and the radiance reflected from the object. This

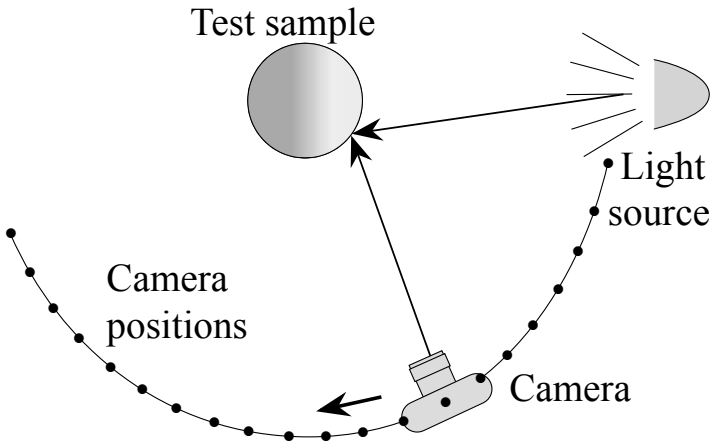


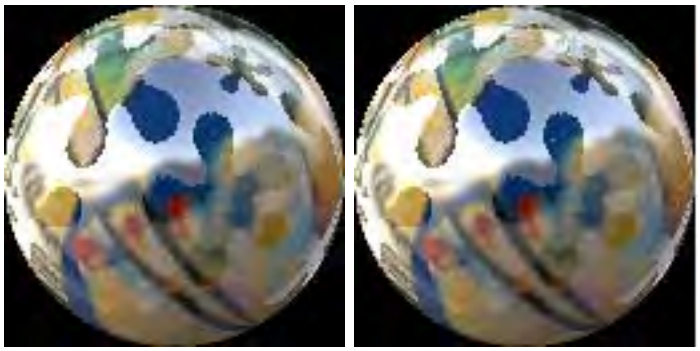
Figure 3.4: Acquisition system used by Marschner *et al.* [50].

Figure 3.5: Objects rendered with estimated BRDFs (Marschner *et al.* [50]).

information is obtained through the calibration of the optical response of the camera and the use of color cards that are used as reference.

This way, it is possible to determine both camera and light directions for each pixel and its surface normal vector. In conjunction with the radiance measured by the calibrated camera, these values enable the estimation of the BRDF of the surface by inverse rendering. The results are used to estimate parameters for the Lafortune model [37]. The work presents some objects rendered with the calculated BRDFs (Figure 3.5), but it does not show color images of the original objects.

Shen and Takemura [64] observed that the specular reflection varies less than the diffuse reflection on the surface of objects. They proposed a method that initially calculates the specular image of the view through the elimination of the diffuse coefficient using low-order spherical harmonics. The specular image is used to calculate the parameters of



(a) (b)

Figure 3.6: Sphere (Shen and Takemura [64]): (a) original image; (b) rendered image.

the Lafortune model [37], based on the minimization of the difference between the values in the specular images and the values estimated by the model.

The function is initially estimated considering only a single material across the surface of the object, and a further subdivision into clusters is made, according to a modification of the specular exponent of the Lafortune model. The surface is subdivided until a threshold is reached, and after that the specular pixels are reorganized according to the created groups. To calculate the full BRDF, the specular value of the pixel is subtracted from the observed value, resulting in a diffuse value. The difference between this value and the value estimated for each pixel using the diffuse component of the reflectance model is minimized, thereby obtaining the diffuse reflectance parameters.

The method assumes a point light source and known illumination and shows good results, calculating with precision the reflection coefficients of the objects. However, the paper does not present experiments with real photographs, only with synthetic objects (Figure 3.6).

Beyond estimating the diffuse and specular reflection, Vogiatzis *et al.* [70] propose a method that also estimates illumination geometry and properties using frontier points found in a set of photographs taken with stereo vision. Frontier points are points on the object surface where the epipolar plane is tangent to the surface of the object (Figure 3.7). In frontier points, the normal vector coincides with the normal vector of the epipolar plane.

In the work, surfaces are considered uniform and non Lambertian. The work assumes the Ward reflectance model [74] that, according to the authors, has a good tradeoff be-

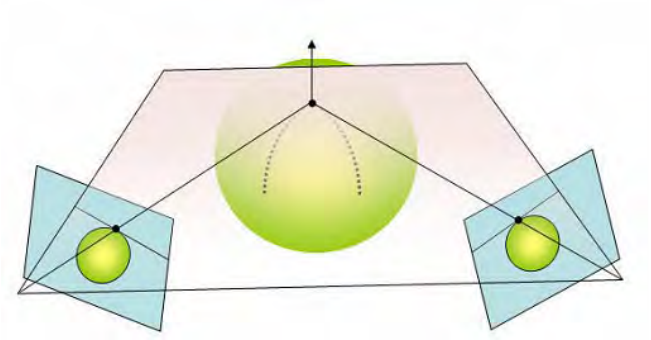
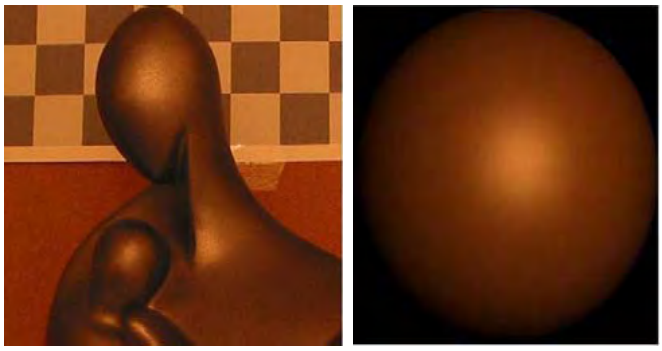


Figure 3.7: Frontier point (Vogiatzis *et al.* [70]).



(a) (b)

Figure 3.8: Sculpture (Vogiatzis *et al.* [70]): (esq) image; (dir) image of rendered sphere with estimated reflection properties.

tween accuracy and computational complexity. This model, together with the equation that represents the irradiance at one point, defines one equation that is minimized in two stages.

In the first stage, the lighting is fixed and the viewpoint is optimized. In the second stage, the lighting is optimized and the viewpoint is fixed. Each stage simplifies the equation to be minimized, and different estimation techniques are used in each one. The minimization is done from the border points of the object (where normal values are known). Thus, the reflectance and illumination values are calculated, and the geometry of the object is retrieved. The work does not present real models with the resulting reflectance, only synthetic spheres (see Figure 3.8).

Lagger and Fua [38] propose a method that also uses the captured images to locate light sources, in an approach that is applicable to objects whose color varies along their surface. It relies on the fact that the influence of the light over pixel values tend to

glide over the surface of the object according to its normal vector, while the influence of the texture over pixel values remain in the same surface position in different images. The work distinguishes maxima lighting and texture on objects and applies a scheme to detect the light source based on the coherence between images. Thus, given a set of images containing a moving object, the method retrieves lighting parameters based on the specular behavior, making less restrictive assertions.

The method assumes that the object texture can be arbitrary, and multiple light sources with unknown colors can be used. It is required that specular reflections appear in some of the recorded images and that the geometry is known. The specular detector is designed to handle light sources that occupy small areas. The method initially detects regions with maximum local intensity, later classified as texture or specular maxima. The pixels in the specular areas are used to estimate the number of light sources and their directions through the use of votes in a Hough accumulator.

Based on the estimated lighting parameters, the direction and intensity of the light sources are refined through the approximation of these values to the Ward model [74] in grayscale, where the 3D model projected onto each image provides information about the normal vectors. The values of the light source are refined by minimizing the difference between the real intensities in gray level and the ones predicted by the model. The reflectance is obtained by interactively minimizing the Ward model. The method defines initial specular and albedo values, fixes the specular values and minimizes the model. Then the specular values are released and the entire system is minimized. Penalties are used to avoid incorrect values. Figure 3.9 shows the method results for a teapot.

Most works on image-based appearance preservation, including those presented in this section, do not consider inter-reflections when calculating the reflectance of an object. Machida and Yokoya [48, 49] present a method to estimate reflectance properties of spatially variant objects that present inter-reflections. In this method, a laser scanner is used to reconstruct the geometry of the object, and multiple images of the object are taken under different illuminations. The geometry and color images are aligned with each other.

(a) (b)

Figure 3.9: Teapot (Lagger and Fua [38]): (a) photograph; (b) rendered model.

The authors use a grid of sixty known positions for the light source and estimate the smallest subset of the best positions before capturing images. The selected positions are the ones that present optimal results for the decomposition of the diffuse and specular components at each surface point. This approach has the advantage that a minimum number of images with different illumination are used. At the end of this process, there should be a subset of positions where virtually all pixels are observed once for the diffuse reflection component and twice for strong specular reflection component.

When this optimal subset is chosen, a single light source is placed at each selected position. The use of a single light source exempts the calibration between the brightness of multiple light sources. The pixels characterized as diffuse are used to measure the diffuse reflection surface using a radiosity equation. This equation assumes inter-reflections between the surfaces and does not consider the ambient light, since the images are captured under a single light source. The specular reflection and surface roughness parameters are obtained through the Torrance-Sparrow model [67], using specular reflection values extracted from the pair of images with strong specular reflection component and the previously calculated diffuse reflection parameter.

The method is experimented in real objects (Figure 3.10), and compared to a previous method developed by the authors. The comparison shows that the consideration of inter-reflections improved the results. However, the authors stress that inter-reflections caused by specular reflectance are not considered, so the results can still be improved.

(a) (b)

Figure 3.10: Ceramic cat (Machida and Yokoya [48, 49]): (a) photograph; (b) image of the rendered model.

3.2 HDR Images

This section describes works that use HDR images (Section 2.4) to model the appearance of real objects. Using this kind of imaging, Matusik *et al.* [51] represent BRDFs of surfaces from a dense set of measurements. The authors argue that the direct use of real information instead of calculating parameters for analytical models enables optimal representations, which are not influenced by local minima or the initial estimates of parameters.

Data is captured at a high resolution and dense angular sampling, thereby obtaining a dense set of BRDF samples for spherical objects. For that, HDR images are captured and a mechanical arm is used to take the pictures in a controlled environment. The authors have chosen a set of materials that could represent a large number of characteristics. To this end, more than 130 kinds of materials were sampled, including metals, plastics and painted surfaces. The resulting set of BRDFs is known as the MERL (Mitsubishi Electric Research Laboratories) BRDF database (Figure 3.11).

The proposed model for isotropic BRDFs is based on the measured reflectance of this wide range of materials. Each acquired BRDF is treated as a vector of high dimension, belonging to the space of all possible BRDFs. Tools are then applied to reduce the dimensionality of this space and find a representation with lower dimensionality. Thus, it is possible to extrapolate and interpolate the space of acquired BRDFs.

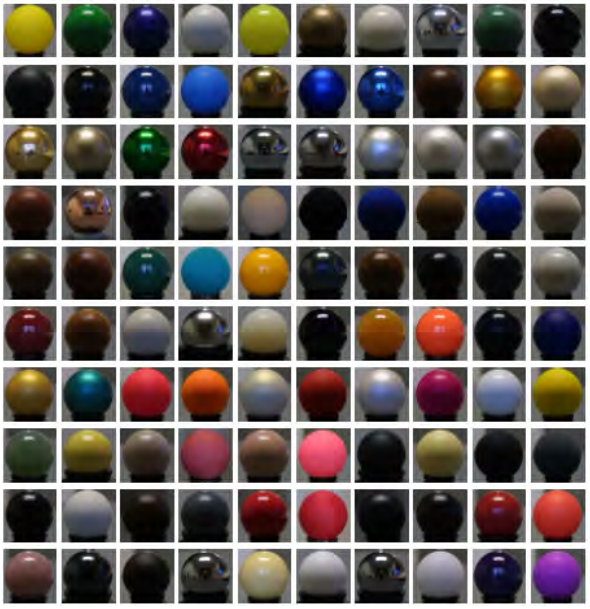


Figure 3.11: One hundred sampled materials (Matusik *et al.* [51]).

Sets of parameters can be defined for navigation in this space based on the perception of the material, which allows the user to create new BRDFs with desired characteristics. The system also allows the user to specify such parameters.

Yu *et al.* [77] present a process to minimize the error during the BRDF estimation that ensures global minima. It models the reflection through the Cook-Torrance model [14] and uses data from the MERL BRDF database [51] as input. Using this database, spheres are rendered with different properties, and the illumination and camera position values are known. The goal is to estimate the parameters of the Cook-Torrance model, which are the Fresnel factor, the diffuse albedo and surface roughness. The authors argue that the Fresnel factor (function of the refraction index and the angle of the incident light) is almost constant at angles of incident light up to 70 degrees. Thus, the work assumes that this value is constant on the surface of the object, and angles greater than 60 degrees are excluded.

Based on the observation that the roughness parameter makes the system nonlinear, a method is described for global optimization which estimates an initial value for the roughness. This value is considered to be the best so far, and a search is made for the best roughness parameter using the Branch and Bound method. This method successively subdivides the range of possible values and analyzes the center of the ranges using a

viability function defined by the authors. This function eliminates the ranges that do not improve the existing values. The subdivision process is repeated until the range reaches a minimum size.

Using the computed surface roughness parameter, the other parameters of the model are calculated using a convex solver. To extract the value of the reflection coefficients, it is necessary to find the intensity of the light source. The results obtained are compared with the MERL BRDF database reference values, and the results are evaluated with the Cook-Torrance model with one and two specular peaks. A comparison is made with the Levenberg-Marquardt, and the results obtained by the proposed method were superior. However, the method is not tested in uncontrolled situations, where imprecise values can affect the results.

Lensch *et al.* [41, 43, 42] present a pipeline to capture, model and display the appearance of objects. It is based on the idea that a material is composed of other materials, so that the appearance of a 3D model can be represented by a set of BRDFs. HDR images are used and the capture is controlled, so that the position of the illumination and geometry of the object are known. Metallic spheres are used to calculate the position of the light source, and the alignment of the object with the 3D model is done through the object silhouette in the photographs. The light intensity is calibrated by imaging a gray card of known color.

Structures called lumitexels are used to represent different color observations at each point of the object surface. A lumitexel contains geometric and photometric information about one point (*e.g.* the position of the point and its normal vector in global coordinates). A lumitexel is associated with an array of reflectances, captured in different images, and with the direction of the light and the viewer in each image. The information stored in lumitexels is used to reconstruct the BRDFs of the object by minimizing an error function between the observed reflectances and the ones calculated by the BRDF using lumitexel values.

The method can be divided into two stages: clustering and projection. In the clustering step a subset of lumitexels is sampled in order to reduce the processing time. These

lumitexels are associated with a BRDF using the Levenberg-Marquardt algorithm. From the covariance matrix returned by the method, the direction where the BRDF parameters can be divided is estimated. A subdivision of lumitexels into two BRDF groups is then made and all lumitexels are relocated in the new groups. The process is repeated until there are no changes in the clustering of lumitexels. The process of subdividing BRDFs is repeated until they have reached m groups, where m is the number of materials defined by the user.

The representation of an object from a small number of BRDFs leaves the object with artificial appearance. This way, a projection step is performed to achieve realistic results and maintain a low image sampling. A number of basis BRDFs is set, so each sample within a group is defined as a projection of these basis BRDFs. The basis BRDFs are initialized with values defined by the authors, and minimized with respect to the samples of the group. The calculated BRDFs are also used in the calculation of normal maps, which can improve the geometry of the model.

The work presents rendering techniques for the estimated models, and shows results obtained with different materials. The number of necessary HDR images is low (around 25). Each HDR image is generated from 15 photos with different exposure times. The method, when combined with related work of the authors (Lensch *et al.* [41, 44]) reduces the number of required views and optimizes the BRDF sampling by using a method to plan the acquisition. Figure 3.12 displays the results obtained with an object and its rendering.

Weistroffer *et al.* [75] present a technique to perform linear basis decomposition from a sparse and scattered set of measurements. Their main idea is that the estimated basis must be defined according to a secondary basis. This way, assuming that the measurements are disposed into a matrix Z, this matrix is decomposed into three matrices: (1) a primary basis weighting matrix W (containing spatially distributed blending weights); (2) a secondary basis weighting matrix H; and (3) a secondary linear basis.

The secondary linear basis is previously fixed, and the authors explored in the paper the use of radial basis functions and a database of measured isotropic BRDFs from MERL

(a) (b)

Figure 3.12: Angels (Lensch *et al.* [43]): (a) original image; (b) image of the rendered model.

database. The matrices H and W are calculated using an iterative approach. First, the matrix W is initialized and the matrix H is updated while W is fixed. After that, the matrix H is fixed and the matrix W is updated. The updating of H and W is based on the resolution of convex Quadratic Programming problems and the iterative process is repeated until it converges to a local minimum.

This estimation technique supports the definition of confidence weights and constraints on the decomposition, making it adaptable to different scenarios, and their results compare favorably with the ones obtained with another iterative decomposition procedure, the ACLS (Alternating Constrained Least Squares). The method is applied in some datasets acquired by Lensch *et al.* [43]; Figure 3.13 shows one of them.

Goldman *et al.* [24] present a method to compute the shape and BRDF of objects using an approach similar to the one proposed by Lensch *et al.* [43]. It is based on the observation that many materials can be decomposed into combinations of a small number of materials. Thus, it is determined that each pixel is composed of a mixture of pair wise combinations of these materials.

The Ward model [74] is used to represent the BRDF, and the user sets the number of BRDFs to be used as basis. From this information and a set of photos with different lighting (the light sources are calibrated using diffuse and reflexive spheres), the method

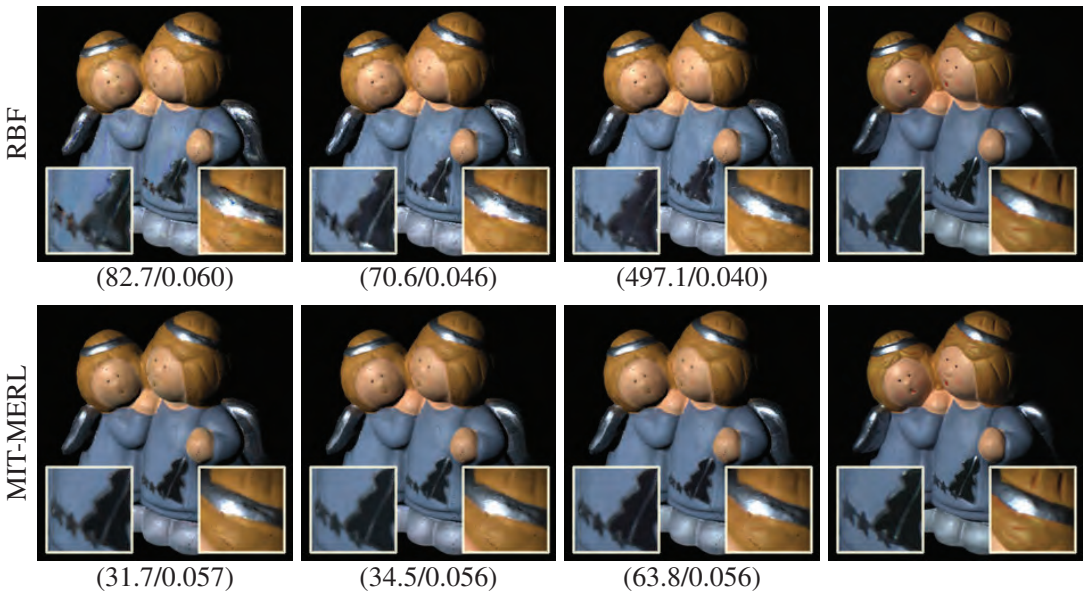


Figure 3.13: Angels dataset (Weistroffer *et al.* [75]): Results obtained by the proposed method using radial basis functions (first row) and MERL BRDFs (second row). The first three columns show results using 64, 1024 and 32768 samples, and the last column shows the original HDR images. Bellow each synthetic image the authors report the running time in minutes and the RMS error.

finds the BRDF of each material, the surface normal and the contribution of each material in each pixel.

The approach to solve the problem alternates between the optimization of the shape given the BRDF, and the optimization of the BRDF that varies spatially due to the shape of the object. First, a normal vector map is initialized using Lambertian photometric stereo, with thresholds to reject specular values. This method also estimates a diffuse albedo, which is used to initialize a weight map. Thus, the normal and weight of each pixel are fixed and BRDF parameters are optimized using least squares. Then the BRDF parameters are fixed and normal vectors and weights of each pixel are optimized together using a scheme that combines brute force search and linear projection. These two steps are repeated until the objective function does not decrease in successive iterations of the algorithm.

The work also shows applications of the method on the edition of BRDFs, and in the transfer of BRDF properties from one object to another. The authors claim that the

(a) (b)

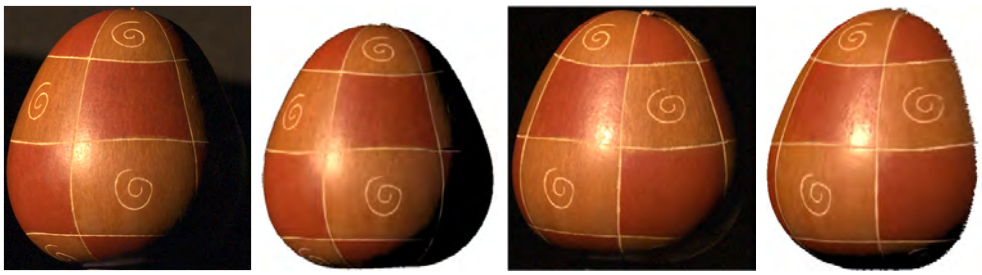
Figure 3.14: Teapot (Goldman *et al.* [24]): (a) original image; (b) image of the rendered model.

method uses a low number of input images, and can reproduce the appearance of different types of materials. Figure 3.14 presents the results obtained with a metal teapot.

Alldrin *et al.* [1] developed a different approach in the calculation of reflectance. In addition to being space variant, the reflectance function is not parametric. Thus, no BRDF model is assumed, and the minimization is done on a set of BRDFs used as a basis, so that the reflectance function at each point is characterized as a linear combination of these basis BRDFs. The number of existing BRDFs in the object is defined by the user.

Each base BRDF is defined by a bivariate function, where the necessary parameters are: (1) the angle between the surface normal and the bisector of the incidence and reflectance directions (*halfway vector* θ_h) and (2) the angle between the incidence direction and the bisector of the incidence and reflectance directions (*difference vector* θ_d). The functions are arranged in a matrix, where its multiplication by a matrix of weights results in the BRDF of the observed point. This formulation defines two sets of parameters: an array B of basis BRDFs and a weight matrix W.

These matrices, together with the elevation angle of the normal $\mathbf{n}$, are obtained through an optimization procedure. This procedure consists of alternating groups of parameters. So, after a random initialization of these values, there is a repetition of a process where B and $\mathbf{n}$ are optimized while W is fixed, and W is optimized while B and $\mathbf{n}$ are fixed. Restrictions are defined for the parameters obtained during the minimization, and the authors found it necessary to define additional restrictions based on the



(a) (b) (c) (d)

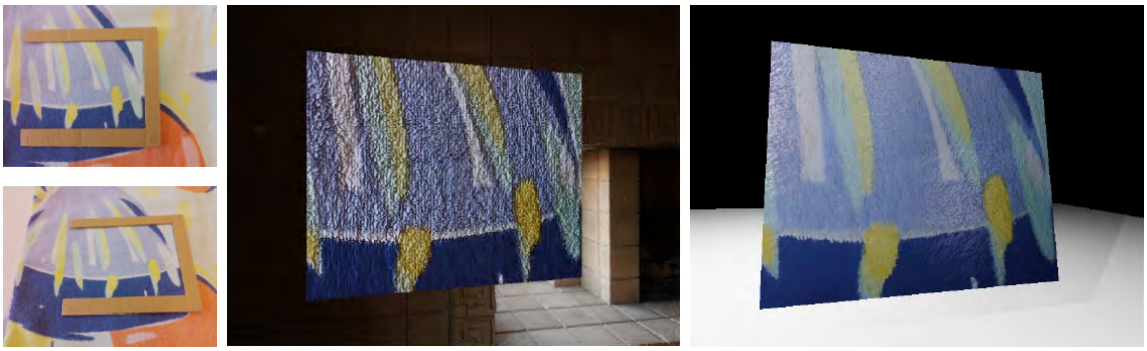
Figure 3.15: Object (Alldrin *et al.* [1]): (a) picture used as input data to the method; (b) image of the rendered object under the viewpoint used to capture (a); (c) picture not used as input data; (d) image of the rendered object under the viewpoint used to capture image (c).

knowledge about the problem domain. Thus, restrictions are imposed in order to assure the smoothness and monotonicity of the BRDFs, and to prevent that specular values dominate the solution.

The work seems to depend on a large set of images with different lighting. In the examples, for one perspective of the object, more than one hundred images with variant illumination have been used. As HDR images are used, it is unclear if these images were used to generate HDR images, or if they already are HDR images (which can considerably increase the total number of captured images). The presented results are very good, and compact representations of reflectance function are obtained (see Figure 3.15).

Häußler [31] estimates the reflectance of planar surfaces using a mobile device camera and a fisheye lens. First, LDR images of a planar surface are captured using a mobile device under different exposure settings in order to retrieve HDR images of the surface. The fisheye lens is connected to a camera, and used to capture the illumination around the surface.

Using this information, the BRDF of the surface is estimated assuming that the BRDF is a linear combination of Ashikmin-Shirley reflectance functions [7]. These functions are obtained through the clustering of the Ashikmin-Shirley functions of a subset of the MERL materials. The environment is aligned to the surface through camera calibration and one environment map is used to represent the incoming light. The normal vectors of the surface are initially estimated through image decomposition. Having defined the scene,



(a) (b) (c)

Figure 3.16: Result from surface normal reconstruction and BRDF estimation (Häußler [31]): (a) pictures used as input; (b) image of the rendered object, relighted with an environment map; (c) image of the rendered object under a point light source.

BRDF estimation approaches are developed for homogeneous and space variant surfaces, and a combined approach that estimates normal vectors and BRDFs is also presented. Figure 3.16 shows results obtained for a piece of fabric using surface normal reconstruction and spatially variant BRDF estimation.

3.3 Color or HDR Images

This section presents a work that can use either color or HDR images as input data. Haber *et al.* [27] estimate simultaneously the reflectance and the lighting of a scene captured under variable lighting and distance. This assertion simplifies capturing and causes the system to be applicable to different scenes using color or HDR images. They assume that the geometry of the object is known or estimated from photographs.

The authors developed a system composed of two main components: first, an all frequency relighting framework is used, based on wavelet representations of: (1) the local visibility; (2) the current illumination estimation; and (3) the reflectance on the point. Thus, it is possible to render the scene efficiently using the integral of the product of the three wavelets. Factoring lighting and scene reflectance requires solving a bilinear system of equations. Thus, the second component employs an interactive optimization in order to estimate the lighting given the reflectance of the scene, and vice versa.

(a) (b)

Figure 3.17: Minerva statue (Haber *et al.* [27]): (a) original image; (b) rendered image.

The paper uses basis BRDFs with the Cook-Torrance model [14]. According to the authors, the recoverable BRDF information is limited by the frequency content of the illumination and geometry of the object, as well as by the environment where it is located. Thus, each independent lighting condition adds restrictions on reconstruction and leads to more accurate approximations of BRDFs.

Assuming that the lighting used for image capture and the BRDF contain a sufficient number of high frequencies, the method is able to reconstruct a detailed model of the scene, lighting and reflection properties (Figure 3.17). According to the authors, a problem that remains is the ambiguity between the color of the lighting and color of the surface. Models generated from images captured in a single environment generate these ambiguities, and this problem can currently be solved by the entry of a sufficient number of photographs of the object in different environments.

3.4 Spectral and Polarised Images

This section presents works that use spectral or polarised images. Li *et al.* [46] present a method for estimating parameters of specular and diffuse reflectance through the use of spectral images. The idea is that this type of image does not have the dynamic range limitation that regular imaging devices possess, so the authors argue that these images are capable of obtaining reflectance parameters with greater correctness.

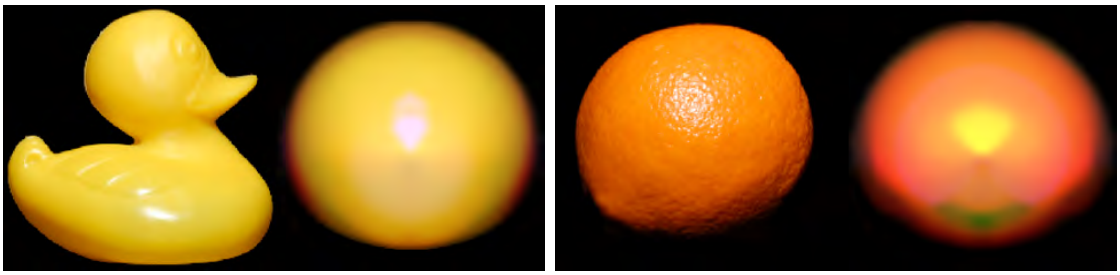
(a) (b)

Figure 3.18: Color cylinder (Li *et al.* [46]): (a) original image; (b) image of the rendered model.

The proposed measurement system comprises an imaging spectrograph and a moving light source, and the spectral images of the object are captured from a sequence of light positions. Reflection values are extracted from the spectral images for each wavelength and are separated into diffuse and specular reflection components. The separation method is based on the physical fact that the intensity of diffuse and specular reflections changes in different ways according to the direction of the incident light.

The reflection values are initially assumed to be the diffuse reflection components. Thus, the diffuse reflection parameter is estimated by minimizing the difference between observed and estimated values at each wavelength, and refined through an interactive algorithm. Then, specular components are calculated by subtracting the estimated diffuse components from the reflection values. A new value for the diffuse reflectance is calculated based on new values of the specular component, and this process is repeated until the difference between the parameters of diffuse reflectance is smaller than a threshold. Using the value calculated from diffuse reflection, the specular reflection component is obtained by minimizing a modification of the Torrance-Sparrow [67] model.

To validate the method, experiments were performed using synthetic values of reflection and actual spectral measurements (Figure 3.18). According to the authors, the method is capable of estimating diffuse and specular reflection with precision, making it suitable for digital preservation of heritage and recognition and classification of objects.



(a) (b)

Figure 3.19: Objects (left) and calculated BRDFs rendered in a sphere (right) (Atkinson and Hancock [8]): (a) plastic duck; (b) orange.

Atkinson and Hancock [8] developed a method that estimates a slice of a BRDF using polarisation data and photometric stereo. The polarisation information (angle and degree of polarisation) is acquired using a linear polariser and a digital camera. This method aims to recover reflectance functions of an unknown material, where its constraints are that the material should have uniform albedo, smooth or slightly roughened geometry and be made of a dielectric material.

Using polarisation the authors retrieve intensity values and an ambiguous estimation of the surface normal vectors. This ambiguity is resolved by a photometric stereo approach presented in the article. The algorithm estimates a part of the BRDF (essentially the ratio of the reflected radiance and surface orientation) by building a histogram of the 3D orientation of the surface and the brightness of the pixels. The method fits a surface that represents the BRDF to the histogram data using optimization by Simulated Annealing [13, 34]. For this, an initial state for the surface and means to disturb this surface are defined.

The method does not work with very rough or metallic surfaces, and retrieves only the diffuse component of the reflectance function. Experiments were done with colored images, which involved the application of the algorithm in each RGB channel (Figure 3.19). However, the color balance obtained from these experiments does not always correspond to the calculated BRDFs. The solution proposed by the authors is to linearly rescale the color components of the BRDFs, so that the color proportions obtained from the BRDFs correspond to the imaged ones.

3.5 Discussion

This chapter presented a review of different image-based works that aim to capture and model BRDFs. As it can be noticed, these approaches are quite diverse, where each one defines a solution according to its context and the type of information that can be obtained from it. Tables 3.1 and 3.2 present the main characteristics of the reviewed methods regarding the considered materials and project design choices, respectively. The remainder of this section is devoted to the discussion of these characteristics.

Table 3.1: Characteristics of the reviewed methods based on the considered materials.

Method	**Appearance**	**Considered Reflectance**
Levoy *et al.* (2000) [45]	Spatially variant	Diffuse
Rushmeier and Bernardini (1999) [59]	Spatially variant	Diffuse
Sato et al. (1997) [61]	Spatially variant	Diffuse and specular
Marschner *et al.* (1999) [50]	Homogeneous	Diffuse and specular
Shen and Takemura (2006) [64]	Spatially variant	Diffuse and specular
Vogiatzis *et al.* (2005) [70]	Homogeneous	Diffuse and specular
Lagger and Fua (2008) [38]	Spatially variant diffuse, homogeneous specular	Diffuse and specular
Machida and Yokoya (2003) [49, 48]	Spatially variant	Diffuse, specular and diffuse inter-reflections
Matusik *et al.* (2003) [51]	Homogeneous	Diffuse and specular
Yu *et al.* (2010) [77]	Homogeneous	Diffuse and specular
Lensch *et al.* (2003) [43]	Spatially variant	Diffuse and specular
Weistroffer *et al.* (2007) [75]	Spatially variant	Diffuse and specular
Goldman *et al.* (2005) [24]	Spatially variant	Diffuse and specular
Alldrin *et al.* (2008) [1]	Spatially variant	Diffuse and specular
Häußler (et al.) (2012) [31]	Homogeneous and spatially variant	Diffuse and specular
Haber *et al.* (2009) [27]	Spatially variant	Diffuse and specular
Li *et al.* (2009) [46]	Spatially variant	Diffuse and specular
Atkinson and Hancock (2008) [8]	Homogeneous	Diffuse and specular

Table 3.2: Characteristics of the reviewed methods based on the project design choices.

Method	**Image**	**Representation**	**Illumination**
Levoy *et al.* (2000) [45]	Color	Average reflectance	Estimated
Rushmeier and Bernardini (1999) [59]	Color	Average reflectance	Estimated
Sato et al. (1997) [61]	Color	Torrance-Sparrow model	Estimated
Marschner *et al.* (1999) [50]	Color	Measured data	Estimated
Shen and Takemura (2006) [64]	Color	Lafortune model	Estimated
Vogiatzis *et al.* (2005) [70]	Color	Ward model	Estimated
Lagger and Fua (2008) [38]	Color	Ward model	Estimated (source not imaged during acquisition)
Machida and Yokoya (2003) [49, 48]	Color	Torrance-Sparrow model	Estimated
Matusik *et al.* (2003) [51]	HDR	Measured data	Estimated
Yu *et al.* (2010) [77]	HDR	Cook-Torrance model	Estimated
Lensch *et al.* (2003) [43]	HDR	Basis BRDFs	Estimated
Weistroffer *et al.* (2007) [75]	HDR	Basis BRDFs	Estimated
Goldman *et al.* (2005) [24]	HDR	Basis BRDFs	Estimated
Alldrin *et al.* (2008) [1]	HDR	Nonparametric BRDF	Estimated
Häußler (et al.) (2012) [31]	HDR	Ashikhmin-Shirley model	Captured
Haber *et al.* (2009) [27]	Color and HDR	Basis BRDFs	Estimated (source not imaged during acquisition)
Li *et al.* (2009) [46]	Spectral	Torrance-Sparrow model	Estimated
Atkinson and Hancock (2008) [8]	Polarised	Measured data	Estimated

Besides the information stored in the images, the reviewed methods can be categorized by the restrictions placed on the characteristics of the material. The main ones are the types of reflections that are considered (diffuse, specular and inter-reflections) and whether or not the materials vary along the surface of the object (see Table 3.1). Most of the reviewed works consider diffuse and specular reflections, with the exception of Rushmeier and Bernardini [59], Levoy *et al.* [45] and Farouk *et al.* [23], whose consider only the diffuse ones. The only method that considers inter-reflections is the one proposed by Machida and Yokoya [48, 49], although only diffuse ones. The assertion of a homogeneous material which does not vary along the surface considerably simplifies the process of calculating the appearance, since all samples at all points of the object surface can be used. When the material is spatially variant, each point (or group of points) can count as an independent BRDF.

Regarding project design choices (see Table 3.2), most of the reviewed works represent the BRDF through analytical models(Subsection 2.3.3.1). This approach enables a smaller amount of stored data and fast rendering. The choice of a particular model is made based on the material to be reproduced and its computational complexity. In this context, Ngan *et al.* [54] present a seminal paper in the field, where they evaluate several analytical models in terms of their ability to reproduce measured BRDFs. Based on numerical errors, function plots, and rendered images they compare the following models: Ward [74]; Blinn-Phong [10]; Cook-Torrance [14]; Lafortune *et al.* [37]; Ashikhmin-Shirley [7]; He *et al.* [28] and a variant of the Ward model proposed by Duer [21]. According to the authors, the ones that presented the lowest errors were the Cook-Torrance, the He *et al.* and the Ashikhmin-Shirley models, followed by the Lafortune model. The Blinn-Phong and the Ward/Ward-Duer models presented the highest errors.

Recent works have been investing in the use of basis BRDFs, where the resulting BRDFs are defined as the weighted sum of a set of BRDFs (*e.g.* Lensch *et al.* [43], Goldman *et al.* [24], Weistroffer *et al.* [75], Haber *et al.* [27] and Häußler [31]). This statement allows a simplification of the problem, such that a smaller number of input images becomes necessary. Three of the reviewed studies (Rushmeier and Bernardini [59],

Levoy *et al.* [45] and Farouk *et al.* [23]) calculate reflectance values at each point in different images, and make a weighted sum of these values to obtain the final reflectance. This approach makes data with extreme values, such as specular reflection values, to be lost. Matusik *et al.* [51] represent a BRDF through a dense sampling of measures, arguing that the direct use of real information instead of calculating parameters enables optimal representations. Marschner *et al.* [50] also represent the BRDF through a set of samples, but the data is fit to the Lafortune model [37] during rendering.

Regarding the illumination requirements (fourth column of Table 3.2), most of the described works assumes a controlled environment, where the positioning of the light source and camera information are known (*e.g.* Rushmeier and Bernardini [59], Levoy *et al.* [45], Sato *et al.* [61], Machida and Yokoya [49] and Matusik *et al.* [51]) or can be retrieved during acquisition. In this context, some works, like Alldrin *et al.* [1], Goldman *et al.* [24] and Lensch *et al.* [43], insert chrome or mirrored balls in the scene in order to infer the position and intensity of the light sources from the reflection observed in these spheres. Others, such as Farouk *et al.* [23] and Marschner *et al.* [50] use patterns and photographic cards arranged in the scene. The methods that calculate the light source in conjunction with the appearance make assertions about the scene to reproduce the appearance of the captured object, such as the need for images of the object inside different environments (Haber *et al.* [27]) and visible specular peaks on the surface of the object (Lagger and Fua [38]). Häußler [31] proposes a different approach, where the environment is sampled using a fisheye lens and used in the BRDF estimation.

It is important to highlight the use of mathematical representations in calculating BRDFs. Haber *et al.* [27] represent the local visibility, the current lighting estimate and the reflectance of the scene using wavelets. Thus, it is possible to render the scene efficiently using the integral of the product of these three wavelets. Shen and Takemura [64] remove the diffuse reflection of images through the use of spherical harmonics, simplifying the calculation of diffuse and specular components of an object. Weistroffer *et al.* [75] use radial basis functions as secondary basis in one of their BRDF representations.

In general, capturing BRDFs requires much time and a large number of images. As it is a function that normally covers four dimensions for each point, planning the capture of BRDFs is not as intuitive as planning the capture of the geometry. Thus, some studies propose methods for estimating optimal positions for the camera and light during the acquisition stage. This makes it possible to uniformly sample the BRDF of objects using a reduced number of images. As described in Section 3.1, Machida and Yokoya [49] choose positions of the light source in order to maximize the diffuse and specular reflection in each pixel of the captured images. Lensch [41, 44] present a system that considers that the camera, the light source, and the geometry of the object are known. Thus, the authors define a function which measures the reduction of uncertainty when adding a new view. This function, along with restrictions imposed by the geometry of the object, is used to predict the next best view of the object. The authors show that the method captures more uniform BRDFs and with less average errors than a human with experience in the subject.

3.6 Conclusions

Based on the review performed in this chapter, it is possible to describe some desirable properties to an image-based appearance preservation method. The data acquisition step is by nature arduous, as many images are necessary to sample the appearance of the surface. This way, the images should be easy to capture and should store reliable information about the reflected radiance. In this problem, HDR images present the best tradeoff, since they store a dynamic range with high resolution and are relatively easy to acquire using cameras that have automatic exposure settings.

When considering the BRDF representation, the use of analytical models reduces the storage requirements and enables faster rendering. The use of linear combinations of basis BRDFs is an approach that enables the estimation of more stable BRDFs, as the BRDF will be defined based on a combination of already existing BRDFs. The basis BRDFs can also be analytical models, resulting in an approach that benefits from both characteristics.

The BRDF models adopted by each approach described in this chapter represent the kind of materials that they aim to preserve. Ideally all kinds of light interaction with matter should be preserved, but this is not a necessary approach in most cases and the number of samples to properly model more complex properties might be unfeasible. This way, a reasonable method should consider spatially varying BRDFs and represent diffuse and specular reflection.

As discussed in Section 1.2, three main approaches can be used to consider the incoming light from the environment during the estimation of the BRDF: to reduce the influence of lighting in order to get a color that is closest to an average; to use direct lighting; and to assume the whole environment as source of light. Considering the environment as source of light provides the more flexible acquisition setup and the more precise information at the cost of more imaging and processing. This way, the choice of the approach should be made according to the purpose of the project, as the complexity of the method increases the more the incoming light is considered.

CHAPTER 4

EXPERIMENTS

In order to understand and apply appearance preservation concepts presented in Chapters 2 and 3, two preliminary groups of experiments were performed. These groups follow the first and second appearance preservation approaches described in Chapter 1: the first group of experiments is made using a method to calculate colors for 3D models by reducing at most the influence of the light that comes from the environment; and the second one aims to capture and represent the reflection properties of a real object using a controlled environment.

It is essential to emphasize the relationship between these works, since the first group of experiments aims at improving the accuracy and correctness of information on the acquisition of colors for object points. The second experiment is a simplified method to capture and represent the appearance of objects with known shape (spheres). This way, to be applied in objects with arbitrary format, it requires a process that captures colors with precision.

This chapter describes these experiments and the works they use as basis. The first experiment is described in Section 4.1 and the second one is detailed in Section 4.2, along with the obtained results. Conclusions about these experiments are presented in Section 4.3.

4.1 First Group: Texture Experiments

These experiments were performed using a texture generation method as basis. This method was developed by the author in previous works and it generates textures for 3D models using color images. This group of experiments has three main purposes: evaluate the method; improve it; and apply the improved method in the preservation of cultural heritage.

A brief description of the texture generation method is made in Subsection 4.1.1, and the evaluation methodology is described in Subsection 4.1.2. The improvements made over the method are detailed in Subsection 4.1.3, and the results obtained with the application of the improved method are shown in Subsection 4.1.4.

4.1.1 Texture Generation Method

Andrade *et al.* [2, 3] developed a method called *Texture*, that generates textures maps for 3D models using high resolution color images. Texture maps are images that contain the diffuse color of the surface of the object, which will be used to render the 3D model using OpenGL's Gouraud Shading [26]. The specular components are manually set to reproduce the appearance of the original object.

Each vertex of the 3D model mesh is mapped to a coordinate at this image (texture coordinate) through a parameterization function. Each triangle in the model is mapped to a triangular region in the texture map, which will contain its texture. To generate the texture map of a 3D object, a simple texture atlas approach is used (Vrubel *et al.* [72]). This approach divides the object in regions with almost planar geometry through the analysis of the vertices normal vectors. Thus, a vertex of the model belongs to a region if the angle between its normal and the average normal of the region is smaller than a threshold, defined according to the object geometry.

The method assumes that the geometry of the object is obtained through the 3D reconstruction of depth images from a laser scanner, using the method proposed by Vrubel *et al.* [72, 71], and that the texture map should be obtained from high resolution color images of the object captured by a photographic camera.

As the geometry and images are captured by different devices, the camera and the scanner should be put in the same coordinate system. To this end, the matrix that performs the transformation between their coordinate systems is calculated using a least squares calibration technique combined with the MSAC estimator [66]. This technique is applied to the set of points of correspondence between the images from the camera and depth images captured by the scanner.

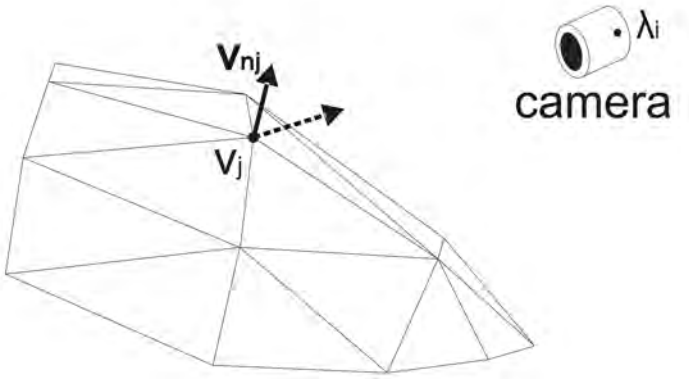


Figure 4.1: Camera perspective of a vertex v_j with normal v_{nj} (Andrade [2]).

By projecting the 3D model over each captured picture, the final texture is generated through the integration of *partial texture maps*. Partial texture maps are texture maps that contain only the color information recorded by a single photograph. As partial texture maps contain many regions with imprecise texture information, each one is associated to a weight map, stored in the alpha channel of the image. This weight map is related to some factors that influences the quality and reliability of the observed color, like the face normal vector direction and the position of the pixel on the object surface during the data acquisition. The partial maps are integrated according to their weights in each pixel, where the result of this integration is the final texture map for the 3D model. Given a partial texture map generated for an image taken by a camera i, the weight w_j of each vertex v_j is initially calculated by the dot product between: the unit vector that goes from v_j to the camera focal point λ; and v_j's normal vector (v_{nj}) (Figure 4.1):

$$w_j = \frac{(\lambda_i - v_j)}{||(\lambda_i - v_j)||} \cdot v_{nj}. \tag{4.1}$$

Where: $\lambda_i = T_i^{-1}\lambda$. The point λ must be in the i coordinate system because being the vertex v_j defined in the coordinate system of the whole 3D model (*i.e.* after global alignment of the range images), the coordinates of the focal point must be in the same reference frame.

Occlusion edges (boundaries formed by parts of the object that appear in front of itself in the images) cause a displacement that is acceptable in most cases, but is perceptible

in occluded regions. In this case, the displacement causes a projection of pixels from the occluding region boundaries into the occluded regions. The areas next to these edges are more susceptible to errors caused by erroneous color projection and integration of the radiance observed in the region, since the colors observed near the edges of occlusion are influenced by two regions that define them (which may have very different colors). This happens due to the configuration of the optical cameras, which calculates the color of each pixel based on the integration of the incident irradiance at the pixel.

This problem is corrected based on the texture map generation method. The correction is applied on the partial texture maps and is based on the comparison between the regions filled with texture in the final and partial texture maps. Considering that the texture map regions correspond to almost planar areas of the 3D model surface, one can assume that when a region is partially filled in the partial texture maps, it is being occluded by another from the camera's perspective. This way, the edges of the final and partial texture map regions are detected. These edges are subtracted and the resulting edges are the boundaries potentially affected by occlusion. The next step is to modify the partial texture map by reducing the weights of the pixels near the occlusion edges. To accomplish this, the occlusion edges are blurred using a Gaussian Filter and the result is subtracted from the partial texture map mask. This way, the reduction is greater when the pixels are near the occlusion edge, gradually decreasing as the distance from the edge increases. When the corrected partial texture maps are merged, the final texture map without occlusion edges is then obtained.

4.1.2 Method Evaluation

Many texture generation methods based on image projection do not present a methodology for texture precision evaluation (*e.g.* Rushmeier and Bernardini [59], Kurazume *et al.* [36], Corsini *et al.* [15] and Dellepiane *et al.* [19]). Also, some of these methods require specific apparatus or assumptions that are not easily reproducible for performance comparisons. This way, a methodology to evaluate the results obtained by texture generation methods is described in this section. Thus, similar evaluation methodologies can be

adopted in future works as a tool for analysis and comparison between different texture generation methods.

In the proposed methodology, three main aspects should be analyzed: the *usability*, the *calibration matrices* and the *accuracy of the final texture map* (*i.e.* if the paintings and visual cues are in the right places). Based on these aspects, the texture generation method used in this experiment is evaluated in this subsection. To this end, 3D models generated from indigenous artworks are used as case study.

4.1.2.1 Usability

Texture presents competitive results with known 2D-3D projection methods, such as Kurazume *et al.* [36] and Dellepiane *et al.* [19]. Kurazume *et al.* project 2D images into 3D models, aligning the edges of the images to edges on the 3D model. Dellepiane *et al.* perform the alignment through the use of correspondence points and presents a graph-based tool to assist the manual selection of points. Their tool infers new matching pairs based on the ones previously found.

The main disadvantage of *Texture* is the need to align the camera and scanner coordinate systems. However, the use of calibration rigs makes this task semi-automatic, and in situations where the rigs may not be used (in the preservation of objects that cannot be moved, for example), the projection matrices can be calculated from automatic or manual search for correspondence points.

4.1.2.2 Projection Matrices

The accuracy of the projection matrices can be evaluated using the 3D-2D correspondence points obtained during the calibration of imaging devices, using the approach proposed by Salvi *et al.* [60]. This approach evaluates the residual distance between: the actual 2D correspondence points (ground truth); and the points obtained through the projection of the 3D correspondence points in the image using the projection matrices (projected 2D points).

Table 4.1: Evaluation of the projection matrices.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Number of pixels per mm	12	10	12	10	8	11
Average residual distance (mm)	0.123	0.093	0.104	0.140	0.204	0.103
Standard deviation of the residual distance (mm)	0.041	0.027	0.040	0.058	0.060	0.043
Maximum residual distance (mm)	0.208	0.160	0.234	0.295	0.336	0.257
Minimum residual distance (mm)	0.022	0.027	0.005	0.015	0.063	0.021
Inliers (number of inliers/total number of points)	(147/147)	(147/147)	(147/147)	(147/147)	(146/147)	(147/147)

Table 4.2: Information about the preserved objects and their 3D models.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Size (cm)	20 × 7.5 × 10	13 × 13 × 10	14 × 5 × 4	13 × 5.5 × 8	26 × 12 × 14	10× 12 × 15
Number of images	54	68	47	52	33	82
Processing time	01:35:30	01:43:05	00:16:22	01:30:19	00:59:25	01:55:11
Texture map size	4096 × 4096	4096 × 4096	2048 × 2048	4096 × 4096	4096 × 4096	4096 × 4096
Number of vertices	72879	160665	81841	37785	64613	92987
Number of faces	131392	294297	154276	67648	117944	170554

As images with different resolutions can be used to generate textures, the average number of pixels per millimeter in the color images should be estimated, and the residual distances converted from pixels to millimeters. Representing these values in millimeters makes it possible to analyze and compare the behavior of projection matrices using different images and methods.

In the evaluation made in this experiment, the number of pixels per millimeter is approximated using two values: the pixel dimensions of the calibration rig in the color images; and the calibration rig dimensions. The division of these values results in an approximation of the number of pixels per millimeter on the surface of the object. Table 4.1 presents an evaluation of the projection matrices used by *Texture*. A description of the evaluated real objects and their 3D models is given in Table 4.2. Although not all correspondence points are considered inliers by MSAC during the calculation of the projection matrices (see line *Inliers*), the results presented in Table 4.1 evaluate all corresponding points obtained during data acquisition, including outliers.

The evaluation shows that the generated calibration matrices present good accuracy, showing very little discrepancy between the projected 2D points and the original ones. The fact the accuracy of the calibration matrix depends on the precision of the correspondence points detection demands special attention. Problematic pictures of the calibration rig (in a too oblique angle or with poor illumination) affect the accuracy of point detection and thus the results of calibration matrices.

4.1.2.3 Texture Maps

In this criterion, the accuracy of the texture map is measured based on the assumption that there is a second texture map generated from images taken using the same optical device that captured the depth images. This assumption spares the need for a calibration step during the generation of texture maps for these images, making these textures a ground truth for comparison. Here, the texture map obtained from camera images will be called *proposed texture map* and the texture map images obtained from the scanner images will be called *reference texture map*.

First, Cumani's method [16, 17] is used to detect color edges in both texture maps. This method was chosen because despite the need for parameter tuning, it is fast (Naik and Murthy [53]) and presents good results (Koschan [35]). Thus, for each edge point in the reference texture map, the nearest one is searched in the proposed texture map. Some edges on the reference texture map may not appear in the proposed texture map due to noise or difference of detail levels. So, a distance threshold was defined; if the point does not have a match within this distance threshold, it is assumed that it does not appear in the proposed texture map. The chosen threshold values were the ones that enables: a low average residual distance; a low residual distance standard deviations; and maximum distance values considerably higher than the average plus the standard deviation.

Table 4.3 shows the evaluation of the texture maps used in the rendering of the 3D models shown in Table 4.2. It is important to stress that the edge detection on the reference texture map can be affected by the low resolution of the images captured by the depth acquisition device, when the algorithm could estimate edges in blurred borders. One example of this problem can be seen in the Model 4 in Table 4.3. As the object texture is naturally blurred, the color images acquired by the depth acquisition device became yet more blurred and its results are not as good as the other objects results. Considering this problem and the high resolution of the proposed texture maps (see Table 4.2), it is possible to assume that an average distance of few pixels is acceptable and that the proposed texture maps are very accurate when compared to the reference texture maps.

Table 4.3: Evaluation of the texture map precision.

	Mod. 1	Mod. 2	Mod. 3	Mod. 4	Mod. 5	Mod. 6
Average distance (pixels)	2.26	1.59	1.75	3.66	2.02	1.78
Standard deviation (pixels)	1.89	1.53	1.54	3.10	1.72	1.50
Maximum distance (pixels)	10	10	10	15	10	10
Minimum distance (pixels)	0	0	0	0	0	0
Distance Limit (pixels)	10	10	10	15	10	10
# points on the borders of the proposed texture map	166393	161560	84309	16237	54787	124328
# points on the borders of the reference texture map	81139	105394	46172	9011	45784	101335
% of points in the reference map that have matches	97.5	97.17	99.21	91.38	96.65	99.3

(a) (b)

Figure 4.2: Ceramic duck (model 4), rendered with textures obtained by different methods: (a) Vrubel *et al.* [71, 72] and (b) Texture [2].

4.1.2.4 Appearance

Comparing the visual appearance of 3D models with textures generated using Texture [2] and a method that uses color images from the laser scanner (Vrubel *et al.* [71, 72]), one can see that the texture resolution has been improved by using Texture, as images taken using a photographic camera provide a higher level of details. Figures 4.2 and 4.3 show comparisons for two of the models evaluated in this chapter (models number 4 and 5, respectively). For more information regarding these models, refer to Tables 4.1, 4.2 and 4.3.

Specular and shadows that appear in a subset of images are reduced because the weights of the partial texture maps consider the position of the pixels in the object during image capture. Figure 4.4 illustrates this property in the 3D model of a ceramic jaguar (model number 6). This property is positive when Lambertian surfaces are considered, as the influence of lighting peaks in the final texture maps is reduced.

(a) (b)

Figure 4.3: Ceramic tapir (model 5), rendered with textures obtained by different methods: (a) Vrubel *et al.* [71, 72] and (b) Texture [2].

(a) (b)

Figure 4.4: Reflectance effect on a Karaja ceramic jaguar (model 6): (a) picture of the object, showing the reflectance effect on its surface; (b) the 3D model texture is not affected by the reflectance effect.

4.1.3 Method Improvement

The *Texture* method (Subsection 4.1.1) was designed to be applied on small and medium-sized objects (*i.e.* up to approximately 50 cm). The method was not tested in larger objects during its development, and as it is presented in Subsection 4.1.4, it was necessary to apply it in bigger objects, with more than two meters high.

The size of the objects influences the size of the generated texture map and increases the computational cost of the method. The larger the map, the greater the amount of memory and processing required. Thus, changes and optimizations were made to the code in order to reduce the processing time and the amount of memory allocated by the method. The method was also decoupled from the system it was part, which allows its execution based on files generated by the geometry generation stage and modularizes the system.

Also, a new criterion was added in the integration of pixels to form the final texture map. As it was described in Section 4.1.1, each point in the final texture map is the result of the weighted sum of all observations of this point on color images captured by the camera. Originally, this weighting takes into consideration the visibility of this point by the camera, and reduces the contribution of observations located in occlusion edges.

The new proposed criterion diminishes the contribution of another kind of edge: the silhouette edges of the object in images. Like the occlusion edges, these edges may contain erroneous color due to alignment errors and influence the irradiance from the background of the original image. Moreover, they are also susceptible to incorrect projections caused by the discretization of geometry: as the geometry of the 3D model is a discretization of the original object, its silhouette can be less smooth than the actual silhouette of the object in the photograph. This can cause the acquisition of incorrect colors in these areas due to the projection of the 3D model on its photographs.

The silhouette edges have modified contribution in each partial texture map. The edges are detected in each image through the projection of the model in the image. Then, these edges are dilated by a mask whose size is proportional to the texture size. The

Table 4.4: Information about the preserved objects and their 3D models.

	Jonas	**Daniel**
Number of images	144	172
Texture map size	8192 × 8192	8192 × 8192
Number of vertices	4288302	3899724
Number of faces	8157634	7336795

affected pixels contribution is then zeroed in the map of weights associated with the partial texture map, so they will not influence the final texture map.

4.1.4 Results

The improvements were made to make the method support a new category of object size and location, aiming objects that are bigger and fixed in a more challenging environment. The improved Texture method was applied in a project developed in partnership with UNESCO, which aims to preserve the statues of the Twelve Prophets (*Doze Profetas*), made by the baroque artist *Aleijadinho*. These statues are carved in soapstone and are located in the Bom Jesus de Matosinhos Sanctuary, in Minas Gerais, Brazil. More information about the place where the objects are located can be found at UNESCO Website [69].

In this experiment, textures maps were generated for some of the statues, each one measuring around two meters high. Table 4.4 shows data regarding the models presented in Figure 4.5. These statues are located outdoors and in places of difficult access, which made the data acquisition process more complex. The first challenges encountered in this project were published in the *Journal of Cultural Heritage* (Andrade *et al.* [5], 2012). Figure 4.5 shows rendered images of the 3D models of two prophets.

The texture generation process and the evaluation method were published in the *Journal of Cultural Heritage* (Andrade *et al.* [4], 2012). An overview of the projects that use the improved *Texture* method was published in a *Special Session on Works in Progress*, which occurred in the *Conference on Graphics, Patterns and Images* (Mendes *et al.* [52], 2010). An article featuring the digital preservation pipeline that uses Texture as texture

(a) (b)

Figure 4.5: Rendered images of the prophets 3D models (Andrade *et al.* [6]): (a) Daniel; and (b) Jonas.

Figure 4.6: Rubber sphere used in the experiments.

generation method was published in the journal *Extensão em Foco*, from *Universidade Federal do Paraná* (Andrade *et al.* [6], 2010).

4.2 Second Group: Capture and Modeling Reflectance Properties of an Object

In this group of experiments, color images of a real object are taken in order to model its reflectance properties through a BRDF. It is important to stress that this method is simplified in various aspects, aiming at three main goals: assimilation of the reflectance concepts studied in this thesis; verification and validation of capture and modeling approaches; and establishment of an initial framework for practical experiments.

In the first stage, presented in Subsection 4.2.1, the environment setup is described and images of a rubber sphere are taken (see Figure 4.6) in order to record the appearance of the sphere material in different positions relative to a light source. In the second stage (Subsection 4.2.2) these images were processed to retrieve spatial data and reflectance samples of the object points. In the third stage, this information was used to estimate the parameters of a BRDF model using an optimization algorithm (Subsection 4.2.3). Finally, the results obtained with the rendering of these parameters and an analysis of the results obtained are presented in Subsection 4.2.4.

(a) (b)

Figure 4.7: BRDF acquisition table: (a) table surface; (b) table cover.

4.2.1 Environment Setup and Image Acquisition

An environment setup was built based on the work developed by Marschner *et al.* [50], described in Subsection 3.1. It is a low-cost interpretation of this method, where a convex object and a light source are placed at fixed positions, and the camera moves along a hemisphere around the object (see Figure 3.4). Thus, it is possible to sample the domain of an isotropic BRDF through a set of two-dimensional slices obtained from each picture. The method assumes a single light source, and that no light is reflected from the environment.

In this group of experiments, a BRDF acquisition table was built. This table is lined with black cotton fabric to minimize inter-reflections caused by the table surface. To reduce the influence of the environment (room walls and furniture) on the color of the object, a cover for the table was assembled. Figure 4.7 shows the BRDF acquisition table and its cover.

Twenty-five pins were fixed along the hemisphere around the object, so that the distance and angle between them is constant. A support was built to hold the camera in a fixed distance from the sphere (see Figure 4.8(a)). This support fits on the pins, so it is possible to put the camera in controlled positions. The sphere is also attached to this support, so it rotates during acquisition and it is always still from the camera perspec-

(a) (b)

Figure 4.8: BRDF acquisition table: (a) camera support; (b) light source position (inside the red square).

tive. This configuration ensures that a certain pixel in two of the images correspond to the same point on the surface of the object.

The lighting consists of an incandescent lamp, positioned at the same angle as the first pin of the hemisphere (see Figure 4.8(b)). According to Mondenard [18] *apud* ISO 14524 [33], to be considered a punctual light source, the lamp must be at a distance of at least twenty times its diameter. The lamp diameter is 6 cm, and its distance from the object during the data acquisition was 150 cm. The distance between the camera and the object was 90 cm.

To capture the images, the camera was placed on all the pins of the acquisition table, except for the pin that marks the position of the lighting (if the camera was placed on that pin, it would be in front of the light source). However, not all images were used to calculate the parameters of reflectance because when the camera is placed at opposite or nearly opposite positions relative to the light source, images with undesirable lighting are generated (see Figure 4.9). Thus, from the 24 possible camera positions, the 20 first positions were used to capture the input images.

For each position of the camera images, five different exposure times were captured in order to choose the image sequence whose exposure time best represents the appearance of the ball along the camera movement. Thus, all images used in the next step of the experiment (Subsection 4.2.2) were captured under the same camera parameters.

Figure 4.9: Effect caused when the object stays between the camera and light source.

Figure 4.10: Edited and reduced image of the sphere.

4.2.2 Image Processing

The images captured during the acquisition stage have a high resolution, which caused imperfections of the sphere to become evident in the images (see Figure 4.6). Thus, to reduce these imperfections, the sphere images were blurred with a Gaussian filter. The images were also resized in order to reduce the number of sampled values. Even with the reduction, 532.930 samples were obtained from 20 images. The image obtained from Figure 4.6 is shown in Figure 4.10 after this preprocessing step.

A spherical object has been chosen so that it is possible to infer its geometry analytically. Thus, it is necessary to process the images to obtain the color, normal vector, the camera direction and the direction of the light source for each pixel belonging to the sphere in the images.

To simplify calculations, it was assumed that the camera position is fixed relative to the sphere position and that the light moves in the images. A virtual coordinate system is then defined, where: the center of the sphere is at its origin; the camera is in the Z-axis;

and that the light rotates around the Y axis in positions on the XZ plane (table). The center of the lamp is at the same hight as the center of the sphere, so the light y value is equal to zero and thus its elevation angle is always zero. This way, the positions of the light source are calculated by rotating the initial position of the light source around the Y axis at known angles (defined by the pins on the acquisition table).

Using simple trigonometric relations, it is possible to calculate 3D point values (x, y, z) for each pixel in the sphere using its radius and center. So, given such a point p, its normalization will result in its normal vector (the center of the sphere is the origin). The normalization of the difference between the position of the light source and p will result in the direction of the light source relative to p, and the normalization of the difference between the camera position and p will result in camera direction relative to p.

Having these values for each valid point (*i.e.* point visible by the camera and by the light source), the next step is to convert them to the space used by the adopted reflectance model (described in Subsection 4.2.3). This space defines that the normal vector should be aligned to the Z axis, and the camera and light source directions should be positioned relative to the new position of the normal vector. For this, rotations around X and Y were calculated to align the normal to the Z axis, and these same rotations were applied to the camera and light source directions, thereby obtaining the vectors $\hat{\mathbf{v}}$ and $\hat{\mathbf{l}}$, respectively.

Using this notation, the normal vector is no longer a required parameter. In this new space, beyond the pair of vectors $\hat{\mathbf{v}}$ and $\hat{\mathbf{l}}$, a point on the object is defined by its BRDF. As the BRDF is a function that depends on the wavelength, it is assumed that different BRDFs are calculated for the R, G and B channels (Lensch *et al.* [43]).

In this work, it is assumed that the BRDF corresponds to the RGB color of each pixel normalized between 0 and 1. This assumption can be justified by using Equation 2.7 to model a BRDF that does not change over the surface. It is assumed that the radiance of the incident light $L_i(\omega_i)$ is constant and equal to 1 at all wavelengths, and that the area of the solid angle formed by the point light source $\mathrm{d}\omega_i$ is also equal to 1. Thus, the division of the radiance of the sample by the dot product between the normal and the light source becomes the observed intensity at each pixel in the R, G and B channels.

The capture and processed images will then result in a set of vectors with the form $[\hat{\mathbf{v}}, \hat{\mathbf{l}}, f_{rR}, f_{rG}, f_{rB}]$, which are used in the estimation of the reflection parameters, described in Subsection 4.2.3.

4.2.3 Reflectance Parameters Estimation

In this stage, the vectors $[\hat{\mathbf{v}}, \hat{\mathbf{l}}, f_{rR}, f_{rG}, f_{rB}]$ are used to estimate the parameters of a BRDF. In this experiment, the isotropic form of the Lafortune model [37] was chosen. For an introduction and evaluation of the main reflectance models vide Weyrich *et al.* [76] and Ngan *et al.* [54], respectively.

The Lafortune model was chosen due to its computational simplicity and generality. The isotropic Lafortune model can be defined as follows (Lensch *et al.* [43]):

$$f(\hat{l}, \hat{v}) = \rho_d + \sum_{i=1}^{m} [C_{xy,i}(l_x v_x + l_y v_y) + C_{z,i} l_z v_z]^{N_i}. \tag{4.2}$$

The model presents a sum of reflection elements, where ρ_d is the diffuse component and a summation is made over m reflectance lobes. Each reflectance lobe i defines a reflection behavior of the model, and depends on the vectors $\hat{\mathbf{v}}$ and $\hat{\mathbf{l}}$ previously described, on the exponent N_i, and the values $C_{xy,i}$ and $C_{z,i}$. N_i represents the specular component of the model, and the higher its value is, the more concentrated the specular reflection will be represented. The signal of $C_{xy,i}$ makes lobe i retro-reflective (positive sign) or not (negative sign). The magnitude of the parameters $C_{xy,i}$ and $C_{z,i}$ defines the albedo of the lobe, and the division of these values represents the reflection in higher incidence angles (off-specularity).

In this study tests were performed with $i = 1$ and $i = 2$, where the set of parameters is calculated for each RGB channel, totaling 12 parameters for $i = 1$ (four per channel), and 21 parameters for $i = 2$ (7 per channel). To estimate the parameters $\rho_{d,k}$, $N_{i,k}$, $C_{xy,i,k}$ and $C_{z,i,k}$ using the measured data in each channel k, the Levenberg-Marquardt optimization algorithm [47] was used. This algorithm requires an initial set of values for the parameters. It has been observed empirically in this experiment that the initial value

Table 4.5: Results obtained with the Lafortune model optimization using one lobe.

k	$\rho_{d,k}$	$C_{xy,1,k}$	$C_{z,1,k}$	$N_{1,k}$	**RMSE**
Red	0,1098588	-0,9038196	0,985324	7	0,0886
Green	0,2505658	-0,1148486	0,495456	1	0,1066
Blue	0,4271027	-0,0517101	0,55003	1	0,1037

Table 4.6: Results obtained with the Lafortune model optimization using two lobes.

k	$\rho_{d,k}$	$C_{xy,1,k}$	$C_{z,1,k}$	$N_{1,k}$	$C_{xy,2,k}$	$C_{z,2,k}$	$N_{2,k}$	**RMSE**
Red	0,0640267	-0,3195667	0,620565	2	-0,9778732	0,96785	15	0,0736
Green	0,2395167	-0,0310047	0,444568	1	-0,9623528	0,931067	15	0,0823
Blue	0,4152349	0,01081284	0,534952	1	-0,9397031	0,839312	13	0,0942

of $N_{i,k}$ influences the convergence of the other model parameters. That is, if the same value of $N_{i,k}$ is used with different values of $C_{xy,i,k}$ and $C_{z,i,k}$ in different executions of the optimization algorithm, $N_{i,k}$ remains the same while $C_{xy,i,k}$ and $C_{z,i,k}$ always converge to the same values in these executions.

In order to avoid local minima, tests were performed where the initial estimate of the parameters varies in $N_{i,k}$, using initial values ranging from 0 to 20. The other parameters were initiated with values from a known BRDF ($C_{xy,i,k} = -0.8$ and $C_{z,i,k} = 0.8$). The best results for the model with 1 and 2 lobules are displayed and evaluated in Subsection 4.2.4. The selection of these results was based on the Root Mean Squared Error (RMSE) between the sampled intensity values and the ones obtained by the Lafortune model using the estimated parameters.

4.2.4 Results

Representations of the Lafortune model using one and two lobes ($i = 1$ and $i = 2$) were experimented, where parameters were calculated for each of the RGB channels. The parameters that achieved the best results in different iterations of the Levenberg-Marquardt algorithm are presented in Tables 4.5 and 4.6, along with the Root Mean Squared Error between the sampled and estimated intensity values.

The representation with two lobules presented lower errors. As it can be seen in Figure 4.10, this happens due to the type of reflectance of the sphere, that is better ap-

proximated by two lobes: one lobe with more concentrated specular reflection and another with a scattered reflection. This is confirmed when the exponents N_i are analyzed in Tables 4.5 and 4.6. They show that the first lobe in both tables features a spread specular peak (lower $N_{1,k}$ values), whilst the second lobe presents a specular peak concentration (higher $N_{2,k}$ values).

Analyzing the error values, it can be noticed that the Lafortune model with two lobes better represented the measured data. It can also be observed that some channels were better approximated than others. This happens due to the fact that each channel presents different reflectance values, and some are better represented by the Lafortune model. The Lafortune model generates values based on cosine functions, so channels that show reflections that behave like this function are best represented.

The estimated parameters for the model with one and two lobes were used on the rendering of spheres under illumination that mimics the ones used during data acquisition (*i.e.* with the same positioning of the light source relative to the sphere). Figure 4.11 shows two types of rendering: one with the values obtained directly from the model, and another with these values multiplied by the scalar product between the direction of the light source and the normal vector of the points, in order to simulate shading. The points on the sphere that are not seen by the camera or by the light source are rendered with black values.

To demonstrate that the renderings are consistent with the values captured by the camera, Figures 4.13 and 4.14 show plots of the intensities of the pixels located in sections of photographs and rendered images. These sections are made in the sphere center (see example in Figure 4.12) to represent the behavior of the surface reflection over a full hemisphere.

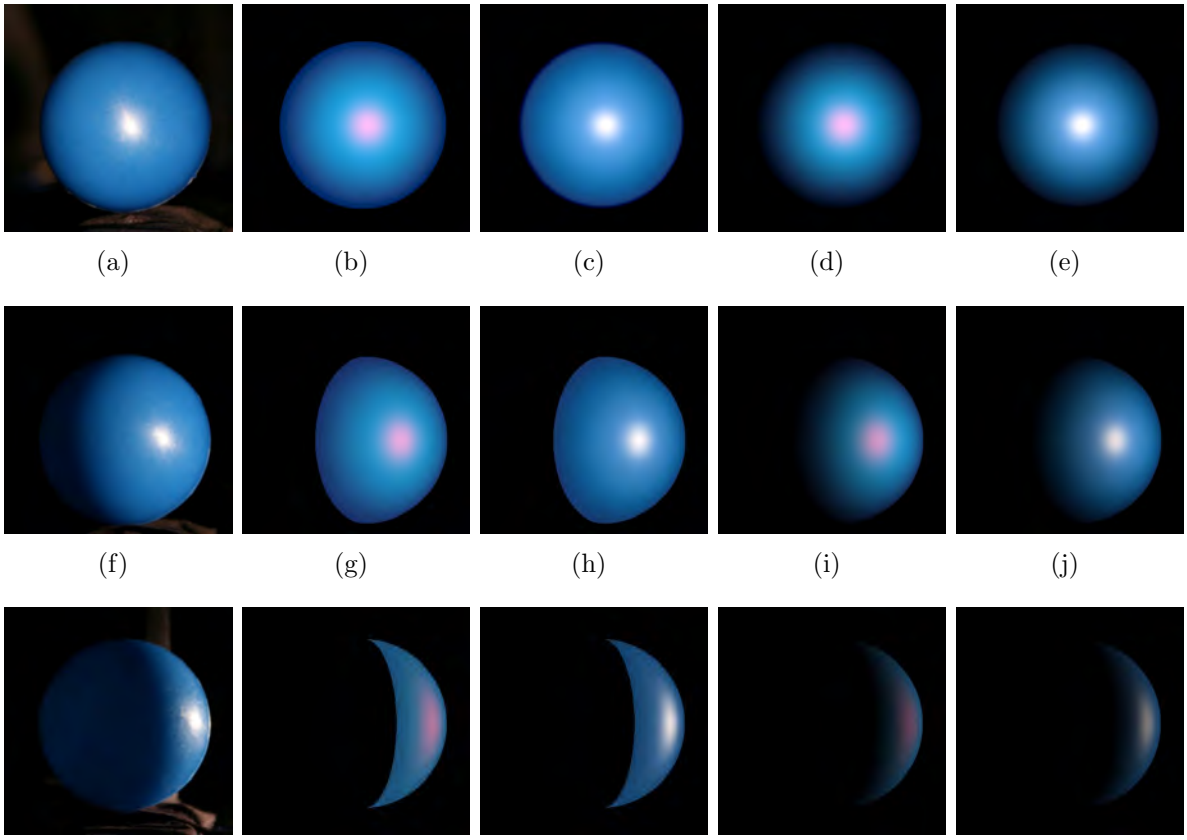


Figure 4.11: Rendered images of the sphere with different positions of the light source: (first column) photographs of the sphere; (second column) rendered images of the sphere using the model with one lobe; (third column) rendered images of the sphere using the model with two lobes; (fourth column) rendered images of the sphere using the model with one lobe and shading; (fifth column) rendered images of the sphere using the model with two lobes and shading.

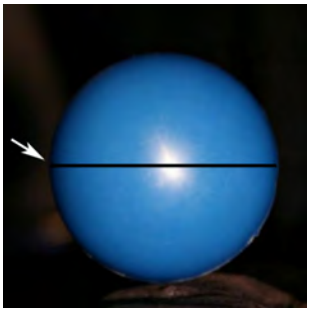


Figure 4.12: Section used in the plotting depicted in Figure 4.13.

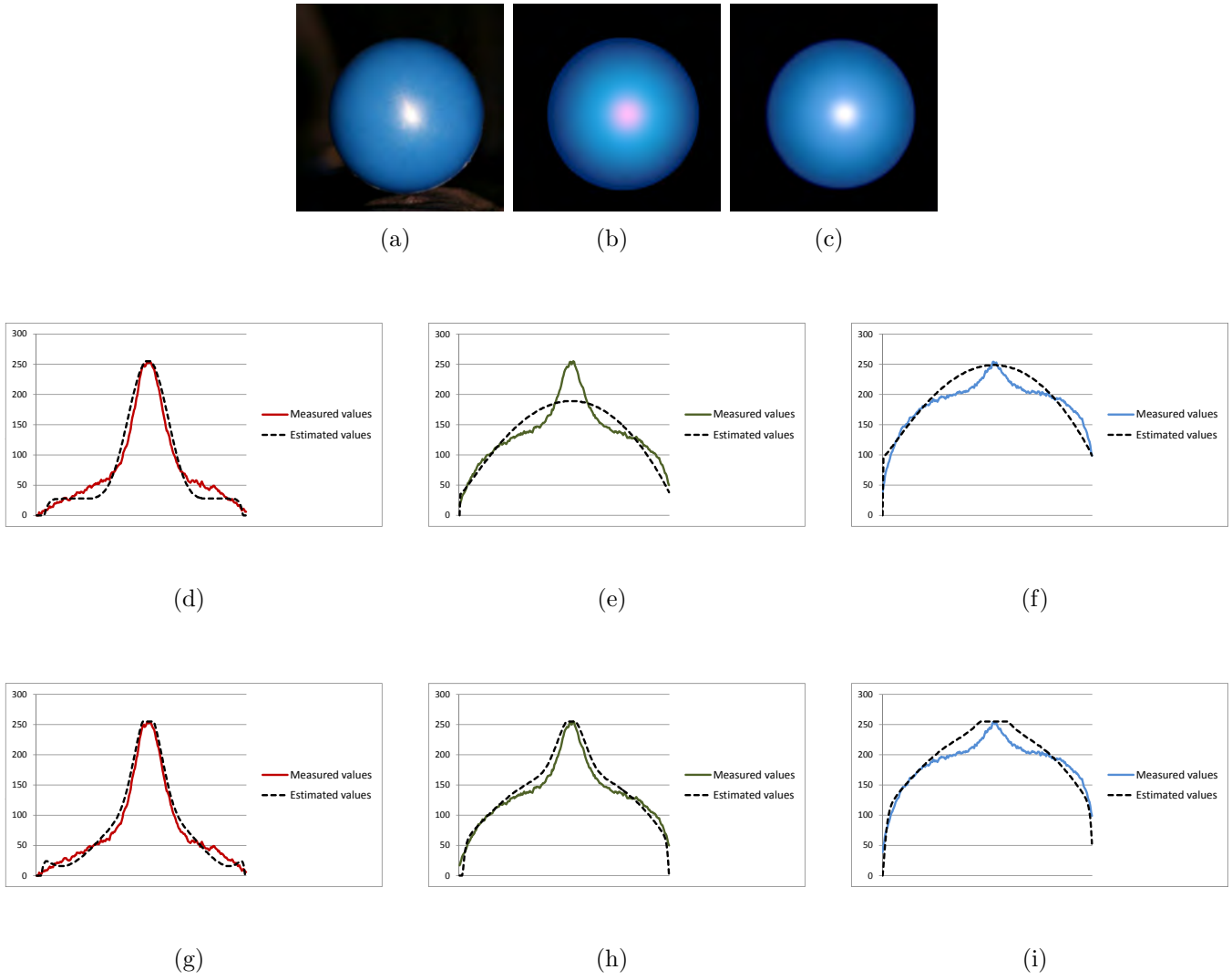


Figure 4.13: Analysis of the observed intensities: (a) photograph of the sphere; (b) rendered image using one lobe; (c) rendered image using two lobes; (d), (e) e (f) comparison between the intensities observed in (a) and (b) in the R, G and B channels; (g), (h) e (i) comparison between the intensities observed in (a) and (c) in the R, G and B channels.

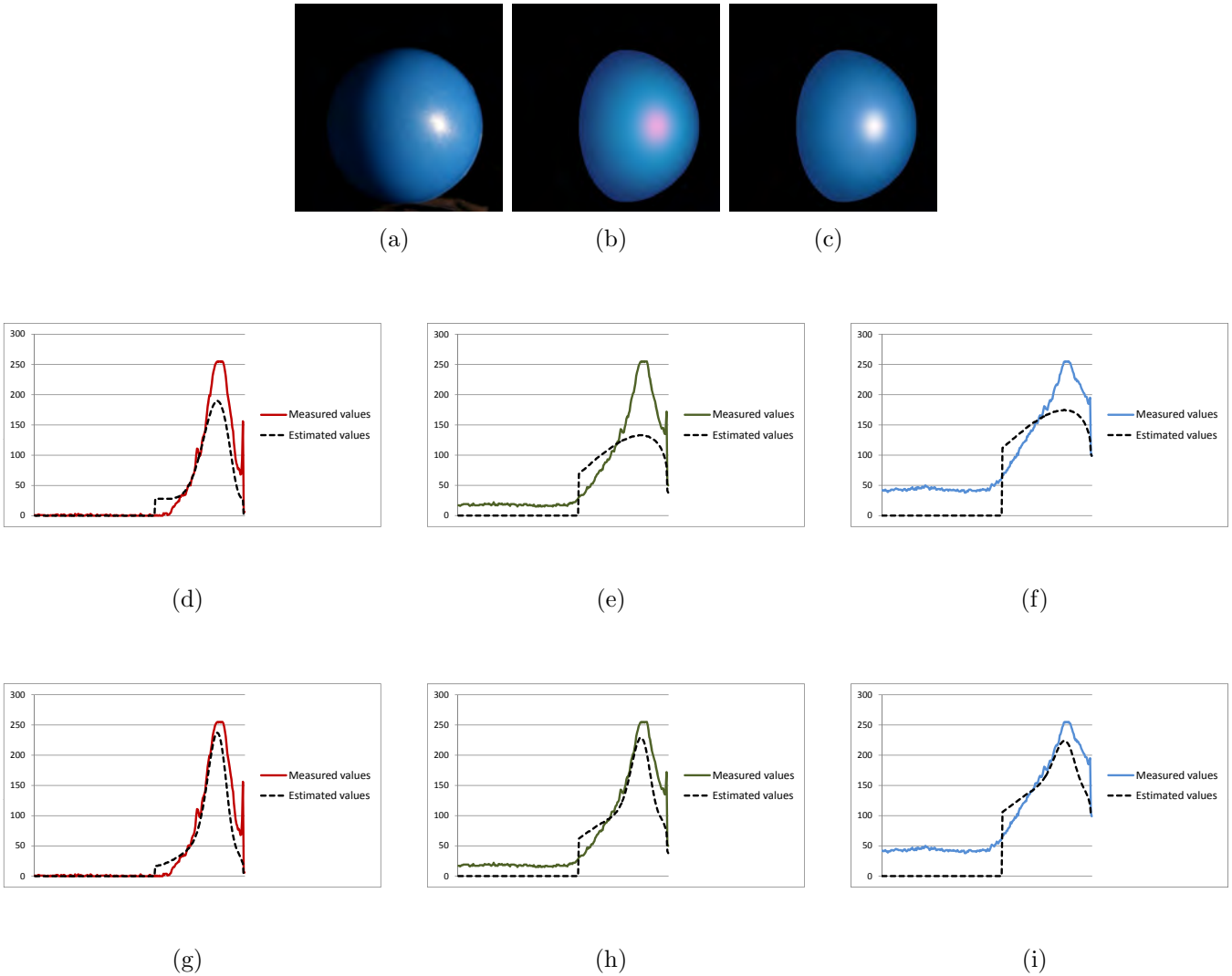


Figure 4.14: Analysis of the observed intensities: (a) photograph of the sphere; (b) rendered image using one lobe; (c) rendered image using two lobes; (d), (e) e (f) comparison between the intensities observed in (a) and (b) in the R, G and B channels; (g), (h) e (i) comparison between the intensities observed in (a) and (c) in the R, G and B channels.

It can be observed that the renderings are very close to the actual values, and that smaller error values represent more realistic renderings of the model. The parameters of the Lafortune model with one lobe do not represent well high and scattered specular lobes. This becomes visible when the sphere is rendered in positions where the specular peak in one channel has more contribution than the others.

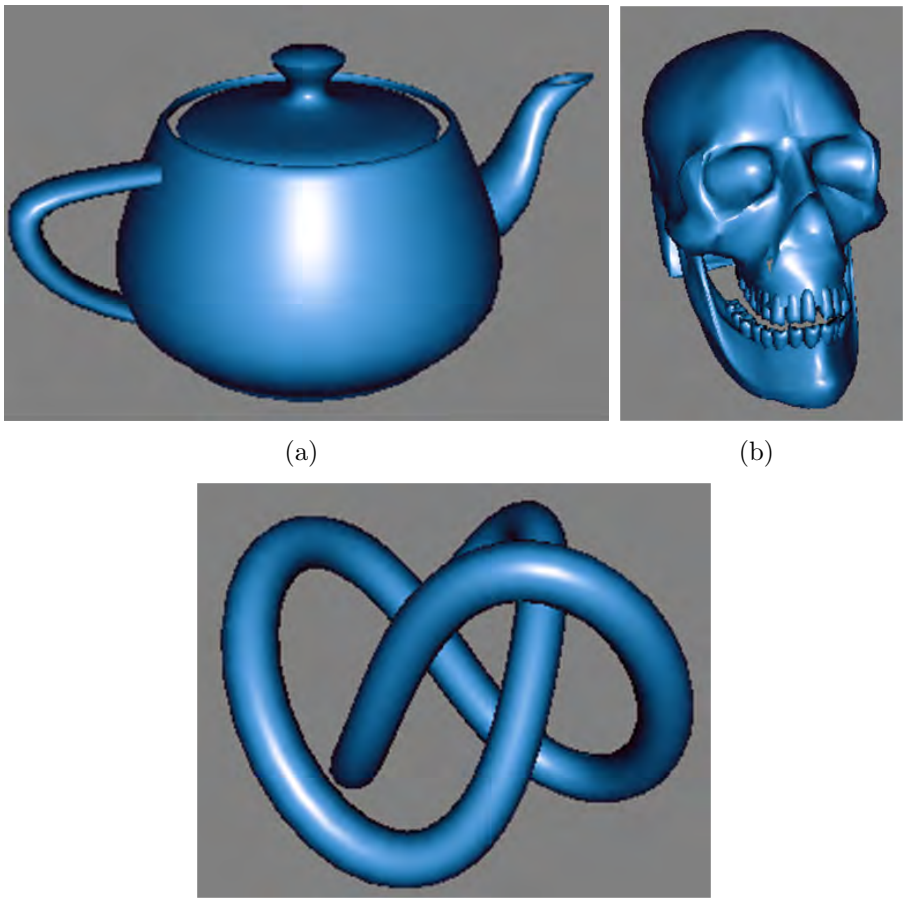
It is possible to illustrate this problem through the plots shown in Figures 4.14(d), (e) e (f). While the red channel has a more concentrated specular lobe (better approximated by the model parameters), the other two channels have more scattered ones, which were approximated by a curve with greater amplitude that "flattened" the specular peaks. This way, as the red specular peak was modeled with an intensity that is higher than the other two channels in the specular peak area, the specular values show a reddish tint in the renderings (Figure 4.14(b)).

The model with two lobes allows a better approximation of real curves, as can be seen in the plots (g), (h) and (i) of Figures 4.13 and 4.14. In this case, in each channel, two curves are combined, one with a more concentrated specular lobe, and another with a more scattered one. Thus, it is possible to reproduce the behavior of real data more precisely and to retrieve more realistic renderings of the sphere. Figure 4.15 shows different models rendered using the Lafortune model with two lobes estimated in this experiment.

4.3 Conclusions

This chapter presented two sets of experiments, made in order to assimilate and apply the appearance preservation concepts presented in Chapters 2 and 3. Each experiment presented contributions, as they use combinations of existing methods and extrapolate them in different aspects.

The first experiment is based on an existing texture generation method. In the experiment, this method is evaluated, improved and applied on the digital preservation of cultural heritage. This way, besides its technical contributions, this experiment also presents cultural relevance. The second experiment is a low-cost interpretation of an



(a) (b)

(c)

Figure 4.15: Renderings of arbitrary 3D models using the estimated Lafortune model with two lobes: (a) teapot; (b) skull; (c) knot. These models were rendered using TyphoonLabs' OpenGL Shader Designer [68]

existing appearance preservation method, and it presents a detailed analysis about the behavior of the chosen reflectance model and about the obtained results.

Each experiment focuses on a different approach on the preservation of the appearance. The first one tries to minimize the influence of the environment in the final reflectance model. The second goes one step further and considers a single light source in this process, which adds more precise information about the behavior of the material under incident lighting.

These experiments were the knowledge basis and also the motivation for the development of an image based acquisition method that considers the whole environment as source of reflectance, presented in Chapter 5. Such method allows more flexibility in the acquisition environments, and considers the incoming lighting from the whole scene, which potentially gives more reliable information about the incident lighting.

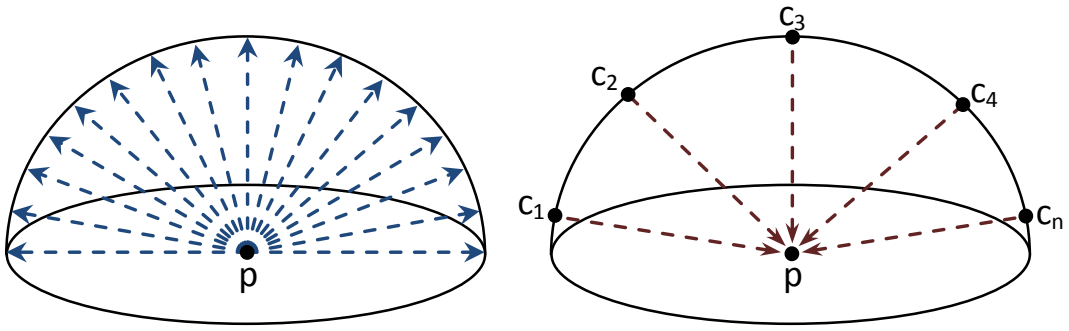
CHAPTER 5

AN IMAGE-BASED PRESERVATION METHOD

During the experiments made in this thesis, it was observed that considering the influence of the incoming light from the whole environment potentially provides more precise results. This can be stated by the fact that in real scenarios the viewed light is the result of an integration of the light that arrives from the whole environment and is reflected by the observed surface. As discussed in Section 1.2, works that consider only light sources as illumination dismiss the influence of all the points that are not light sources, losing an important contribution to the behavior of the material appearance.

The approach of reducing the contribution of the environment by subtracting it or neutralizing it is also valid, but it limits the acquisition and excludes the environment lighting contribution. Considering these observations, this chapter proposes an image-based method that considers the incoming light from a static environment to each point in a surface patch. This lighting information, together with images of this patch from a number of camera positions, are used to estimate the BRDF of each point.

The main problem that arises from this approach is that two sets of images would be necessary to estimate the BRDF: a dense number of images of the scene and a dense number of images of the point. To capture this data, different camera observations of each point are necessary, and samples of the incident lighting from the environment to each point must be obtained. If the images are taken aiming at a single point and at its incoming light, the number of samples could be reduced but the data acquisition stage would be restricted to a very small area in the scene. When a surface patch containing a set of points is considered in this scenario, the capture of these images becomes unpractical, since it will be necessary to calculate the incoming lighting for each point and to capture a set of images of the patch from a regular sample of camera positions around it. Also, if



(a) (b)

Figure 5.1: Rays sampled in a vertical section of the hemisphere around a surface point p: (a) rays traced in order to simulate the environment map around p; (b) rays traced in order to simulate n virtual cameras c_n that view p.

the user wants to estimate the BRDF of other surface patches in the same scene, a new set of specific images must be taken for each patch.

The solution proposed in this thesis to make this problem manageable is based on seeing this data as two sets of rays. The first set contains rays that come from a point p to the environment in a regular set of directions inside a hemisphere around the normal vector of p. These rays end in the points where they intersect the scene geometry. The second set contains rays that comes from a regular sample of virtual camera positions around the point and ends at the point.

If it is possible to get the color where a ray ends as seen from the point where it starts, these sets can be respectively seen as descriptions of an environment map for a point and a set of views of this point (see Figure 5.1). Moreover, they contain the basic data needed to estimate the BRDF of this point, assuming its normal vector and position can be retrieved from a geometry reconstruction stage.

To calculate these rays and their colors, geometry information is also necessary to define the rays ends and starts. A method to estimate the colors in the positions pointed by the rays is required and a suitable and robust BRDF estimation method should be used. This chapter describes how different methods and ideas were combined to solve these problems, providing a BRDF estimation method that presents the following properties:

- Considers that only HDR images of the scene are used as input;

- Assumes that the whole environment is source of light;
- Calculates the BRDF per point in a surface area.

It is important to stress that these properties are limited by the information available in HDR images of the scene, as they are the only source of information. The more HDR images of the scene, the more the results tend to correction. Section 5.1 presents an overview of the proposed method. The data acquisition setup used to capture the HDR images is described in Section 5.2 and Section 5.3 shows how the geometry reconstruction is performed. The light rays tracing is presented in Section 5.4, and the way they are used to calculate the environment maps and viewed colors for a point is described in Section 5.5. Section 5.6 shows how this data is used to estimate a BRDF model of each point. The results and evaluation of the proposed approach are presented in Section 5.7 and its final considerations and limitations are presented in Section 5.8. The partial results obtained at each step are shown in their respective sections.

This method was developed in collaboration with Prof. Dr. Hendrik Lensch, from Universität Tübingen. A sandwich internship was made by the author in Germany under his supervision, sponsored by the CAPES PDSE project (process number 0157-12-0).

5.1 Method Overview

The proposed method, illustrated in Figure 5.2, aims to model a reflectance function for each point belonging to a surface patch inside a scene, thus providing a SVBRDF for this patch. To this end, a set of HDR images is taken inside this scene and used to reconstruct its geometry. This geometry will be shown to the user, where he/she will be able to define the patch whose appearance will be preserved. The normal vector and positions of the points in the patch are obtained through ray tracing. Once the patch is defined, an environment color map and a set of colors viewed from regularly sampled positions are estimated for each point in the patch. This is performed by tracing rays inside the scene geometry in the directions described in the previous section, and by estimating their colors.

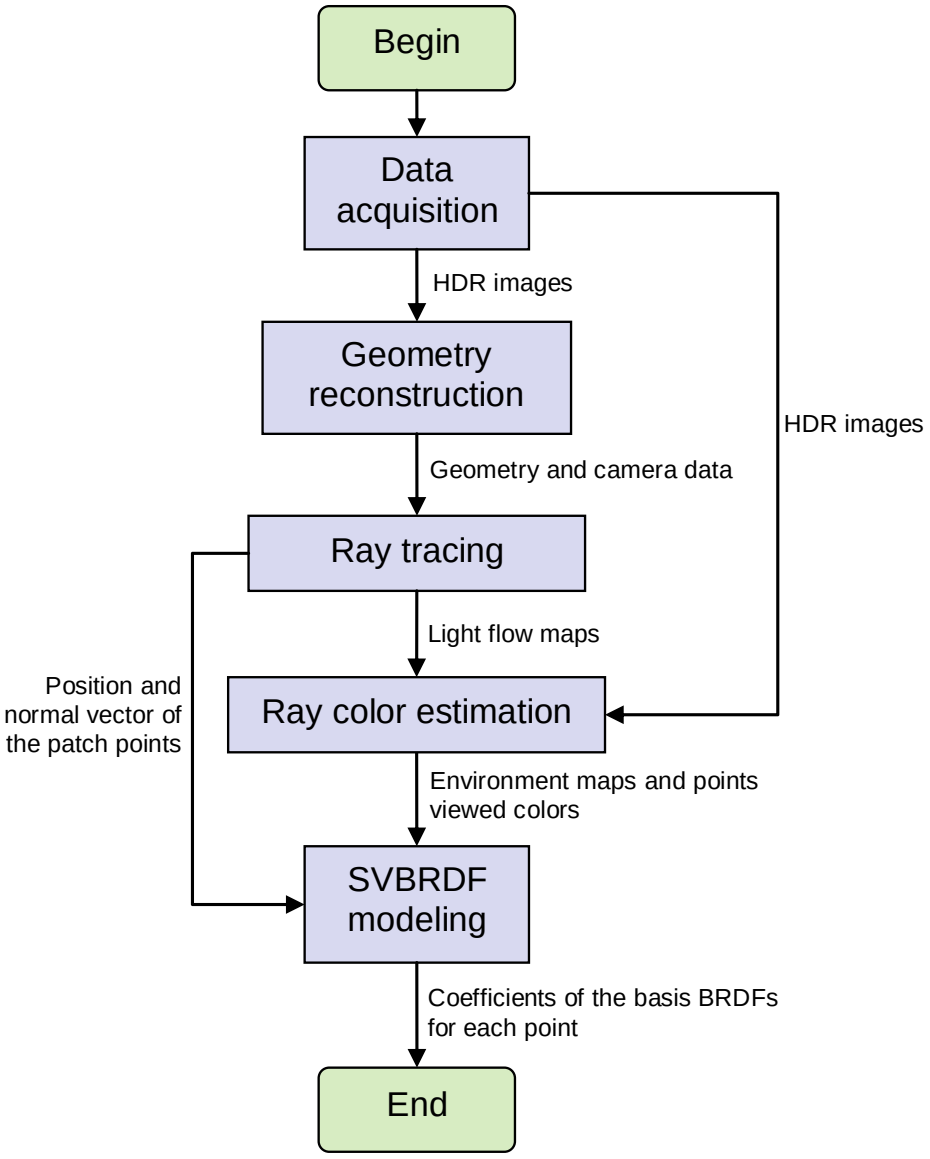


Figure 5.2: Flowchart of the method.

The color estimation for a ray is based on the HDR images taken of the point where the ray ends, in a ray interpolation scheme that uses images of the ray end point as input. Once the environment map and the set of observations of the surface point are estimated, they are used as input to a SVBRDF estimation method, together with data regarding the points positions and normal vectors. The BRDF of each point is defined as the weighted sum of a set of BRDFs. This approach enables a fast BRDF estimation and a greater stability of the results, as the final BRDF will be a linear combination of known BRDF functions. The following sections contain details regarding each step of the proposed method, including the obtained results and performed evaluation. It is important to notice that all the HDR images shown in this chapter were tone mapped using *idisplay*, an HDR image visualization tool, and their brightness was adjusted in order to assure that all details are visible.

5.2 Data Acquisition

During data acquisition, omnidirectional images of the scene were taken using a Ladybug 3 camera (Figure 5.3). This device is composed of six CCD sensors that, according to the Ladybug 3 Technical Reference Manual [58], cover more than 80 percent of a full sphere. For each Ladybug camera position, several images are taken in different exposure settings for each sensor in order to obtain High Dynamic Range images. The images are rectified in order to correct the distortion caused by the lenses.

Several sets of six HDR images are taken, where the Ladybug is positioned in different places inside the scene in order to sample the environment. The number of sets depends on the geometry of the environment. It is desirable that the environment should be regularly sampled in a configuration that approximates a grid. The Ladybug camera was chosen in order to simplify the acquisition approach, as a set of six images is taken from each position of the camera. Nevertheless, it is important to stress that this method also supports HDR images captured by regular cameras.

Figure 5.3: Ladybug 3 camera, from Point Grey Research.

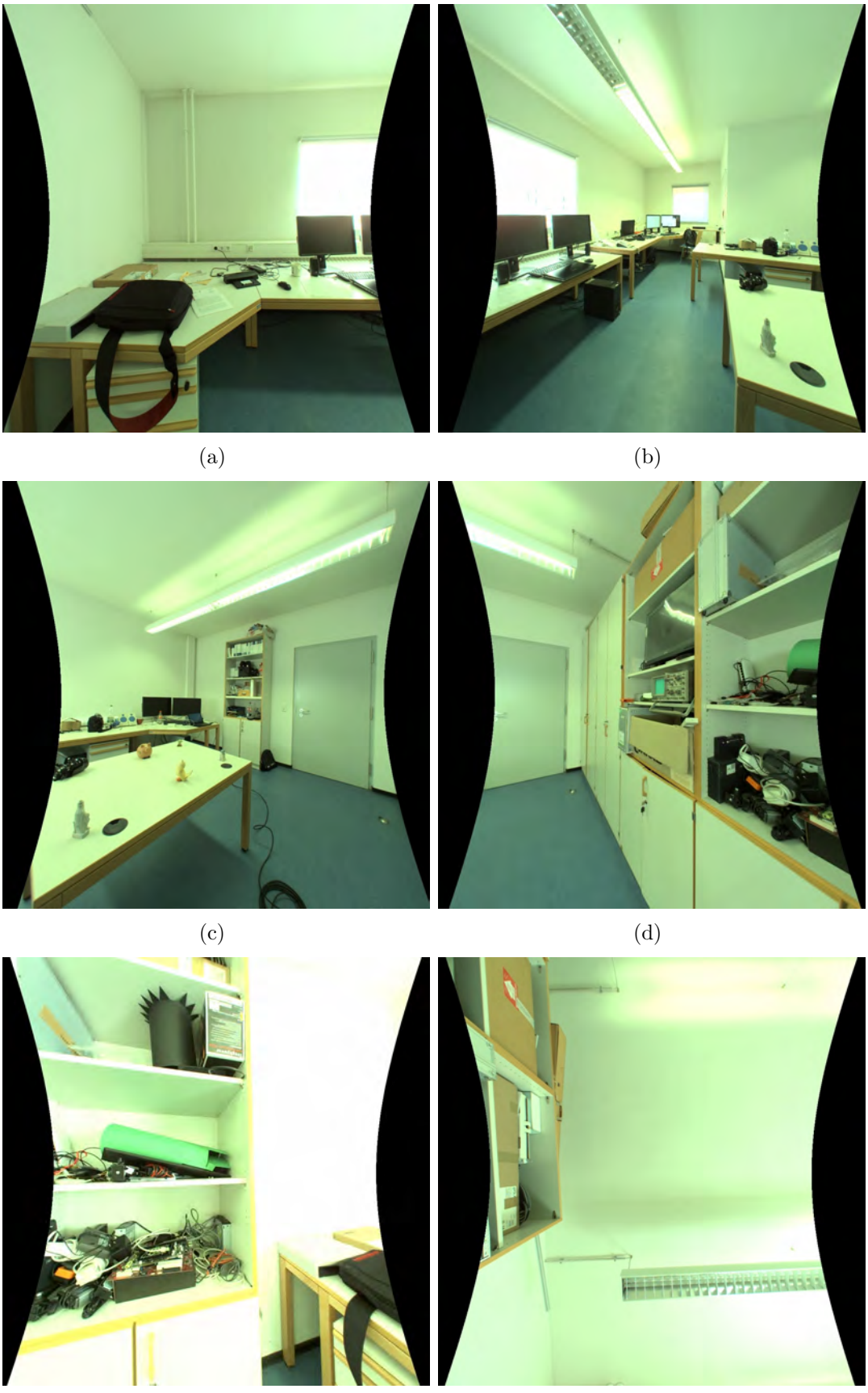
5.2.1 Scenes

Two scenes sampled with the Ladybug camera are used along this chapter to evaluate the proposed approach. The first one is the *office* data set, that depicts an office inside Tübingen University's Sand building. This scene contains natural lighting coming through the windows and fluorescent lights as main light sources. As the walls are white, they are also an important source of light. Figure 5.4 shows a set of six images taken from a Ladybug position in this scene.

The second scene was taken inside Tübingen University's Neue Aula building (*neue_aula* dataset). This scene was imaged at night, so the main sources of light are the lamps positioned around the walls and ceiling, which give to the scene a strong yellow tone. This scene contains a statue inside it and some stairways, as can be seen in Figure 5.5.

5.3 Geometry Reconstruction

The HDR images of a scene are first used to reconstruct its geometry, using combinations of Structure From Motion (SFM) and Multi View Stereo (MVS) methods. Two SFM methods were used: one specific for Ladybug images (developed by Benjamin Resch, from Prof. Dr. Hendrik Lensch research group), and Bundler [65] for images taken using regular cameras. The MVS method was developed by Bailer *et al.* [9].

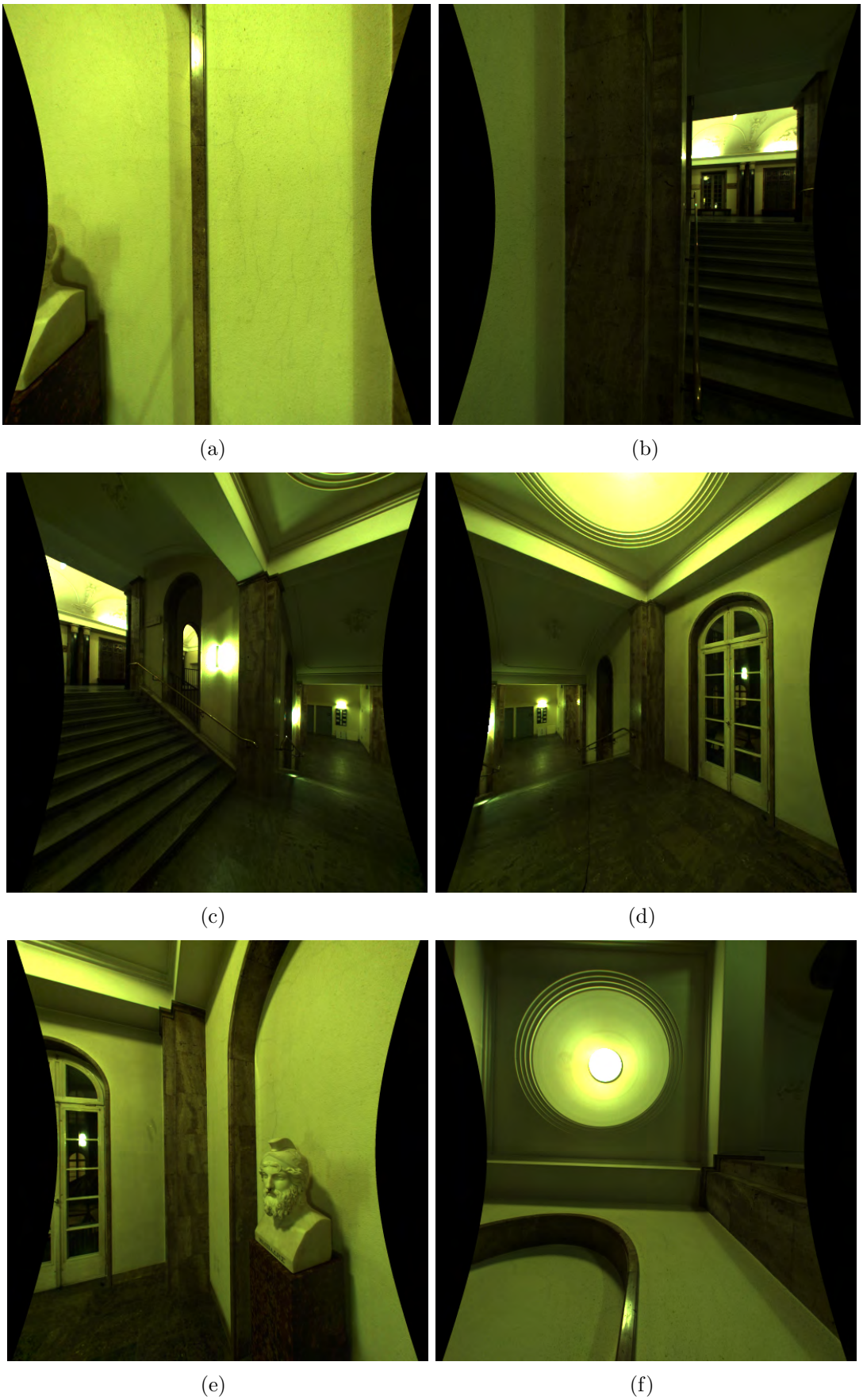


(a) (b)

(c) (d)

(e) (f)

Figure 5.4: Set of six images taken by a Ladybug camera in one position inside a room (*office* dataset).

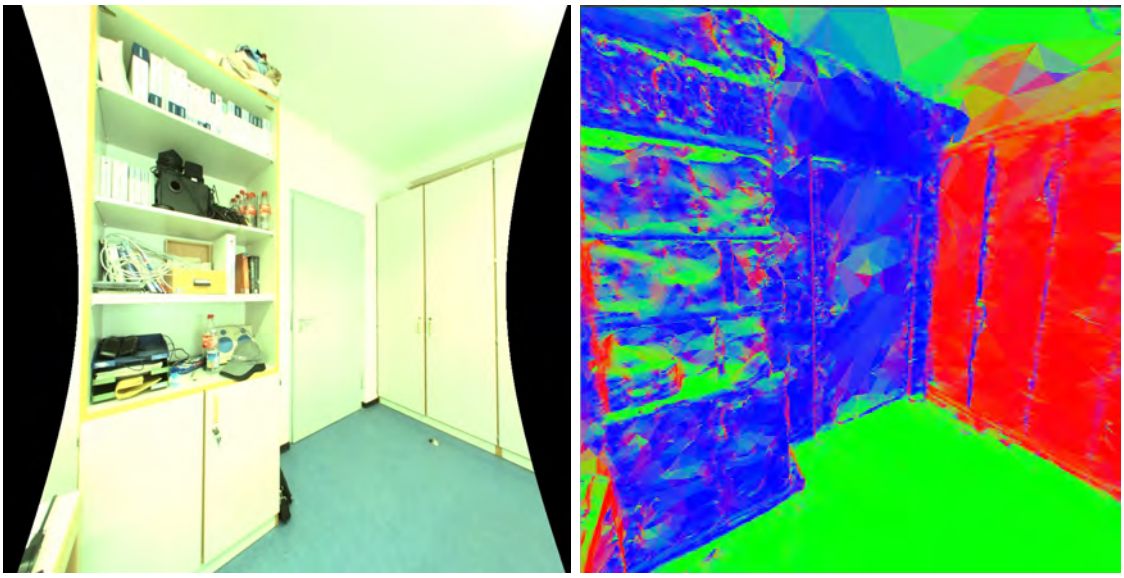


(a) (b)

(c) (d)

(e) (f)

Figure 5.5: Set of six images taken by a Ladybug camera in one position inside Neue Aula building (*neue_aula* dataset).



(a) (b)

Figure 5.6: Geometry reconstruction of *office* dataset: (a) rectified and tone mapped image taken by one of the Ladybug sensors; (b) reconstruction using the Ladybug SFM + MVS approach.

These methods were adapted and integrated in order to retrieve the desired outputs for the next step of the proposed method. This way, given the set of images, the SFM stage returns a set of points in the geometry of the environment and the relative camera positions. This data and the images are then used by the MVS method, which finds out a denser set of points and reconstructs a triangle mesh from them. The MVS method requires LDR images, so all HDR images are tone mapped in order to be used as input to the MVS method. The tone mapping method used was the one developed by Drago *et al.* [22], available in *idisplay* tool.

Figure 5.6 shows one of the Ladybug shots of the office in Figure 5.4 and the 3D reconstruction of this area in the office. The faces in the mesh show color values that correspond to their vertices' normal vectors. It is important to stress that this was a sparsely sampled environment (composed of 180 images of an entire room, with size 1024×1024), so the geometry presents low resolution.

As will be seen along this chapter, a rough estimation of the environment geometry is sufficient for the method's purposes. The only restriction is that the patches that will be sampled for the SVBRDF estimation must present a smooth geometry, without noise

(a) (b)

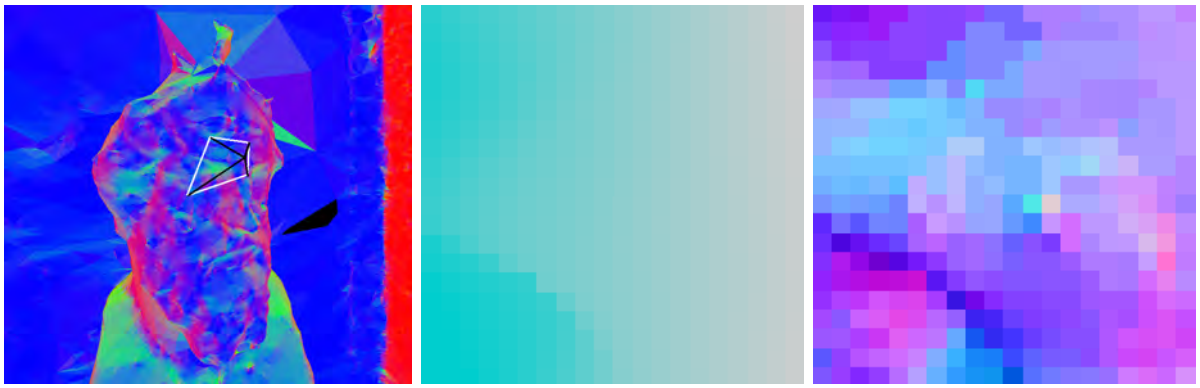
Figure 5.7: Geometry reconstruction of a dataset acquired by Bailer *et al.* [9]: (a) LDR image taken by a Canon EOS5D camera; (b) reconstruction using the Bundler SFM + MVS approach.

or holes. The more images are used, the better are the results obtained. To illustrate this affirmation, figure 5.7 shows pictures of a sofa taken with a Canon EOS5D camera. This is a densely sampled environment with more features (composed of 82 images of the sofa from different perspectives, with size 1853×1236), so the resulting geometry is more precise.

5.4 Ray Tracing

Once the geometry of the scene is known, it is possible to shoot rays inside this scene in order to retrieve the configurations described in Figure 5.1. These configurations will allow the assembling of the light flow of interest, and will be used in Section 5.5 to estimate the environment color map and viewed colors for each point.

The selection of the surface of interest is made by the user through a program that shows him/her the geometry of the scene. In this program, the user can navigate inside the scene in a first person perspective. To select the surface patch to be preserved, he/she must "look" at a surface and press a button to tell the program that this is the surface of interest. Rays are shot from the user virtual eye position to the scene. The area size of the surface is defined by the proximity or distance to the surface patch, and the sampling resolution of the surface is defined by the user. The positions of the points inside the



(a) (b) (c)

Figure 5.8: Results of the ray tracing step regarding the patch (*neue_aula* dataset): (a) surface patch area, sampled by the user; (b) image containing the patch points positions; (c) image containing the patch points normal vectors.

patch are defined through the intersection of rays in the scene. The normal vectors of these points are calculated, as this information is necessary to estimate the SVBRDF of the surface (Section 5.6).

The number of rays is defined by the desired resolution, and they are regularly sampled in order to retrieve two images: a point position image and a normal vector image. These images will delimit the surface patch area and will contain at each pixel the position of the surface points and their normal vectors, respectively. Figure 5.8 shows the result of a surface patch sampling in a scene and the resulting images. These images are composed of float values, being treated the same way as HDR images.

Once the positions and normal vectors of the patch points are known, the light flows for the environment maps and viewed colors can be composed. Subsections 5.4.1 and 5.4.2 describe how they are computed for each point in the surface patch.

5.4.1 Environment Maps

The environment maps aim to represent the light that comes from the environment to the point. This way, assuming that it is possible to get the color of the point where a ray ends as seen from the point where it starts, an environment map can be generated from a set of rays leaving the point into the scene (Figure 5.1(a)).

These rays must be sampled in a way that ensures a nice coverage of the area around the point. The chosen approach is to first sample the rays inside a hemisphere around the point's normal vector by mapping the pixel position (x,y) to polar coordinates (θ,ϕ), as described in Equation 5.1.

$$\begin{aligned} u &= x/(res_x - 1), \\ v &= y/(res_y - 1). \\ \phi &= 2 * \pi * u, \\ \theta &= \frac{\pi}{2} * v. \end{aligned} \tag{5.1}$$

Where (res_x, res_y) is the resolution of the environment map. To improve the sampling, it is assumed that $res_x = 4res_y$. After that, the direction in polar coordinates is converted to cartesian coordinates, projected on a basis where the normal vector of the point is aligned to the z axis of the coordinate frame and stored in the (x,y) position in the image.

This approach also presents the benefit that it makes it possible to retrieve a pixel position in the environment map given a direction in the scene, feature that will be important during the SVBRDF estimation (Section 5.6). Figure 5.9 shows two light direction maps obtained for points in a patch. Each point in a patch has its own environment map, and each pixel in the map contains a direction inside the hemisphere around p.

A ray tracing algorithm is used to define the end point of the rays that come from the patch point in the directions defined by its environment map. These rays are traced along the scene and their hit positions are saved in another image, also using the mapping equations described in Equation 5.1 to define where it will be stored. When a ray leaks through a hole in the scene, the position of its start is stored in the corresponding pixel. This way, the ray color estimation step will be able to recognize rays that never hit a point in the scene surface. Figure 5.10 shows the hit points obtained for the directions presented in Figure 5.9. A subset of the rays used to generate the environment map of

(a)

(b)

Figure 5.9: Environment maps containing directions for points in the patch showed in Figure 5.8: (a) the lower left point; (b) the upper right point.

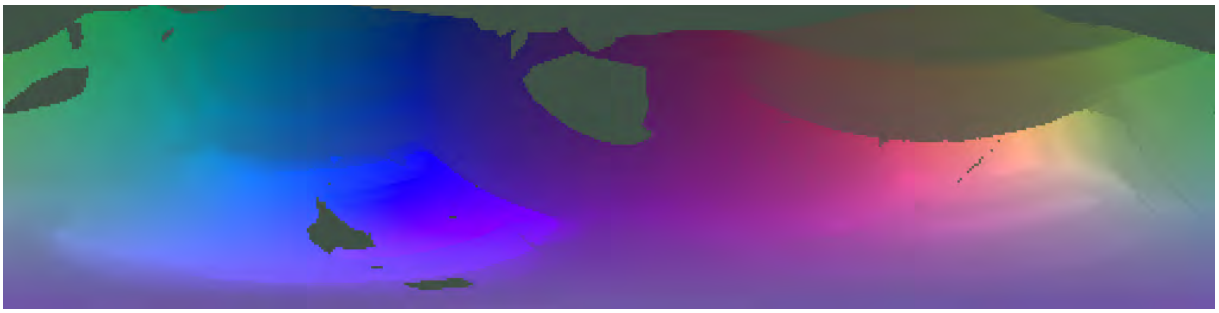
the lower left point in Figure 5.8 is traced in Figure 5.11. These rays are defined by the directions and endpoints contained in Figure 5.10(a).

5.4.2 Viewed Colors

The viewed colors for each point simulate the observation of this point from a number of different perspectives. The idea is to define a set of virtual cameras around the surface patch. Each camera would be the start point of rays that end in each point of the patch. This way, for each point in the patch there will be a ray coming from each camera to this point, as illustrated in Figure 5.1(b). To better illustrate the approach, the results will be shown for each camera, so it will be possible to estimate how each camera sees the surface patch.

This information is necessary in the SVBRDF estimation, as it shows how the color of the point changes according to the viewer position. It influences the specular reflectance values, as a change of colors between observations indicates specular behavior and also

(a)



(b)

Figure 5.10: Hit points found: (a) from the lower left point in Figure 5.8 in the directions defined in Figure 5.9(a); (b) from the upper right point in Figure 5.8 in the directions defined in Figure 5.9(b).

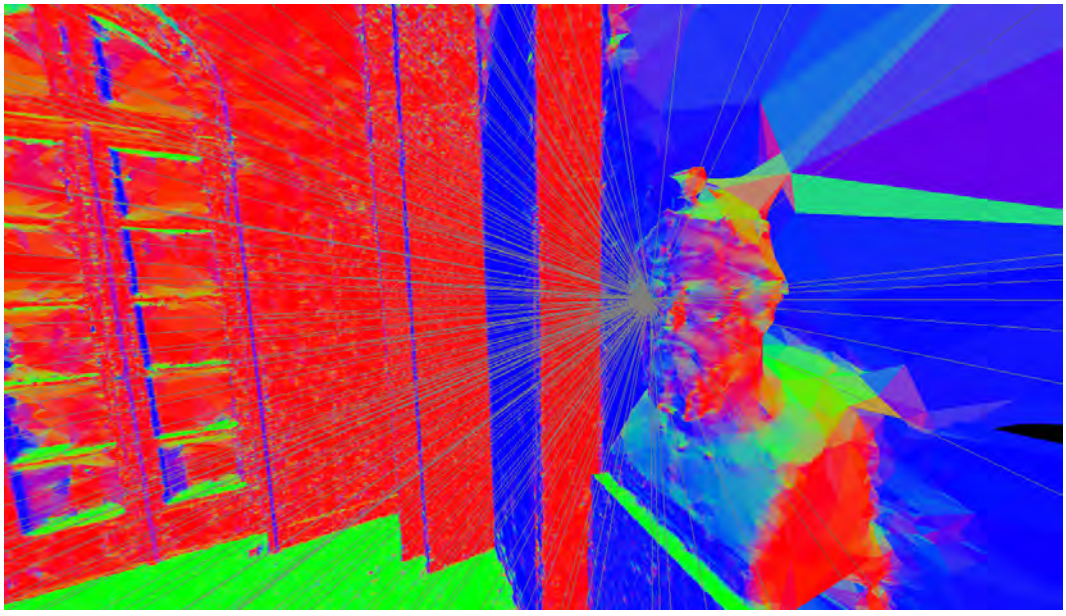


Figure 5.11: Subset of the rays used in the generation of the environment map of a point.

shows the diffuse behavior (*i.e.* the constant light reflected from the surface in all directions).

Color values are obtained by positioning the camera at different (and sometimes pre defined) positions around the interest area, as seen in most works described in Chapter 3. The approach of modeling this problem as a set of rays provides flexibility regarding how this sampling will be made, as virtual cameras can be positioned in any position inside the scene. It was noticed in the experiment made in Section 4.2 that a regular sampling around the interest area provides a nice coverage of the BRDF samples, so this approach was chosen here. The difference is that it is not possible to assume that the surface is a sphere anymore, so the cameras will be positioned in a sphere around the surface patch.

To choose uniform positions of virtual cameras around a sphere, a trigonometry approach has been chosen. The approach is based on an Archimedes' theorem [63], and its application in the regular sphere sampling problem is described in Equation 5.2.

$$\begin{aligned} r &= \sqrt{1 - z^2}, \\ x &= r * \cos(t), \\ y &= r * \sin(t). \end{aligned} \tag{5.2}$$

Where z is uniformly distributed in the range $[-1, 1]$ and t is uniformly distributed in the range $[0, 2 * \pi)$. The difference between samples of z (dz) is defined according to the number of virtual camera positions (n_{cam}). The difference between samples of t (dt) is the golden angle (see Equation 5.3).

$$\begin{aligned} dz &= \frac{2}{n_{cam}}, \\ dt &= \pi * (3 - \sqrt{5}). \end{aligned} \tag{5.3}$$

In the loop to find each virtual camera position (Equation 5.2), z begins with value $(1 - dz/2)$ and is decremented by dz at each iteration; t begins with value 0 and is incremented by dt at each iteration.

This method provides a regular sampling of a sphere with radius of size 1. The virtual camera position values are then mapped to a different radius size. This way, the sphere radius size is defined based on the patch size using some knowledge about how the ray colors will be estimated in the next step (Section 5.5). The colors will be estimated based on the interpolation of real camera data, giving heavier weights to the real cameras that are closer to the ray's start point. It is interesting that the real cameras that are closer to the patch contributes more than real cameras that are far, as more resolution and less scene irradiance interference may be achieved. This way, the virtual camera positions should be as close to the patch as possible.

In this work, the sphere radius is the longest distance between the points located in the opposite corners of the patch. This way, a sphere that involves the whole patch is defined and some margin is given to protuberances. All the rays that represent virtual cameras that view points in the correct side of the patch (defined by the surface points normal vectors) are used as input to the next step of the proposed approach. Figure 5.12 shows an image of the rays defined by virtual cameras around the surface patch depicted in Figure 5.8.

5.5 Ray Color Estimation

The purpose of this stage is to estimate local light fields, *i.e.* given a set of rays inside the environment, to estimate the color of the point where they hit it. Each ray can be defined by its starting and ending points, or by its starting point and direction. In this last case the ending point will be the point where this ray hits the surface, found using a ray tracing algorithm.

So, given a ray, all the images taken of the environment can contribute to estimate the color of the point where it hits the surface. The camera positions and directions are known from the SFM stage, and metrics are used to evaluate which cameras have better

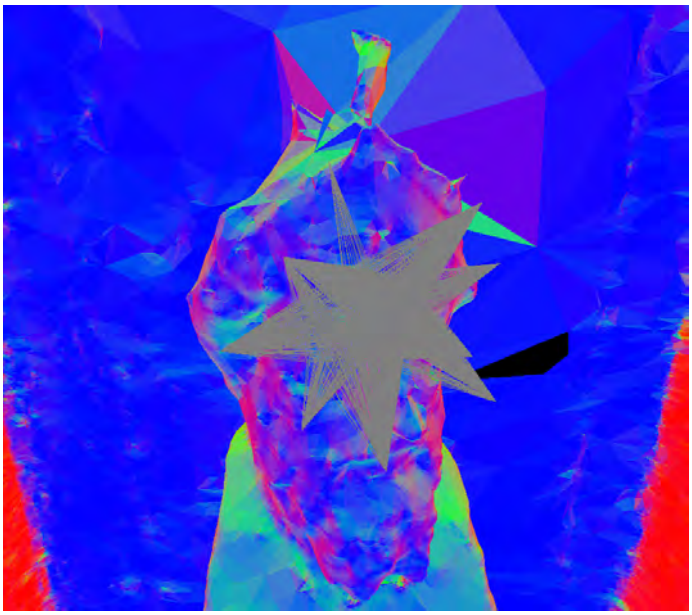


Figure 5.12: Rays traced to simulate virtual cameras around a patch. Only the rays that represent visible points from the camera position to the surface patch are traced

contribution. This way, the final color in the hitting point will be a weighted sum of the colors viewed in that point position by a subset of cameras.

One analogy to this situation is a person who closes her/his eyes inside the environment and points towards a direction. All the cameras will them be used to find what is the color of the point the person is pointing at, based on what they recorded in their images. However, some cameras view the point better than others (and some do not see the point at all), so metrics should be established to find out what are the best ones.

The color at the end of each ray gives the amount of light reflected towards the camera and not towards the point of interest at the start of the ray. Because there is no real camera placed at this point, there is no way of telling the exact intensity of light arriving along that ray direction, unless assumptions are made. This way, the metrics used to choose and weight the best cameras to estimate a desired ray are based in an Image Based Rendering (IBR) method, the Unstructured Lumigraph Rendering (ULR) (Buehler *et al.* [12]). This method is originally applied on the estimation of new pictures of a scene in real time, given an estimation of the scene geometry and a set of images of this scene. As each picture can be seen as a set of rays, the ULR problem is reduced

to two questions: how must the rays be sampled to generate these pictures; and how to retrieve their colors.

In the context of this work, only the second question is of interest, as Section 5.4 describes how the rays are sampled for the specific cases discussed in this chapter. So, regarding the ray color estimation, the ULR method aims to choose k cameras that have the smallest penalties, defined by the:

- *Angular difference:* angle between the ray and the camera direction;
- *Resolution:* camera distance to the point of interest;
- *Field-of-View:* evaluate if the point is inside the field of view of the camera or close to the image borders.

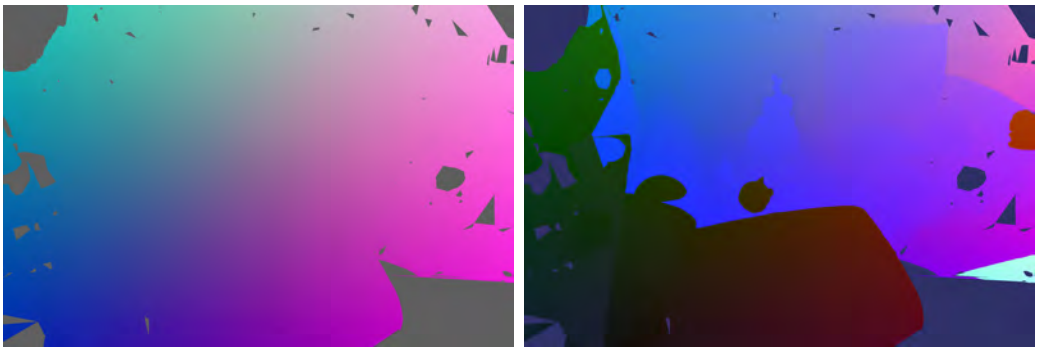
All the cameras receives penalties regarding these metrics, and the k ones with smallest penalty are candidates to contribute in the color. The value k is defined by the user. The candidates are weighted according to an adaptive threshold, defined as the penalty of the $k + 1^{th}$ camera in the penalty ranking ($threshold$); the weight value of this camera is considered to be zero, and the k candidate cameras are weighted as described in Equation 5.4.

$$weight_k = 1 - \frac{penalty_k}{threshold}. \tag{5.4}$$

This way, smaller penalties means higher weights in the color contribution. The rest of the cameras receive a weight equal to zero. These weights are normalized in order to sum to one, and the final color will be the weighted sum of the point colors viewed by the k best cameras. Subsection 5.5.1 describes how the penalty is calculated for each camera.

In this work, sets of rays are stored in two images, given as input to the ray color estimation stage. One image contains the points where the rays start, and the second image contains the points where the rays end. This way, each ray is represented by the values stored at the same pixel position in these images.

In order to illustrate this approach and show its effectiveness, Figure 5.13 shows images containing the direction and ending of rays traced in the environment depicted in



(a) (b)

Figure 5.13: Each ray is defined by the data stored in the same pixel position in different images: (a) direction of the rays; (b) end of the rays. The image containing the beginning of the rays was omitted in this case as all of them start at the same point.

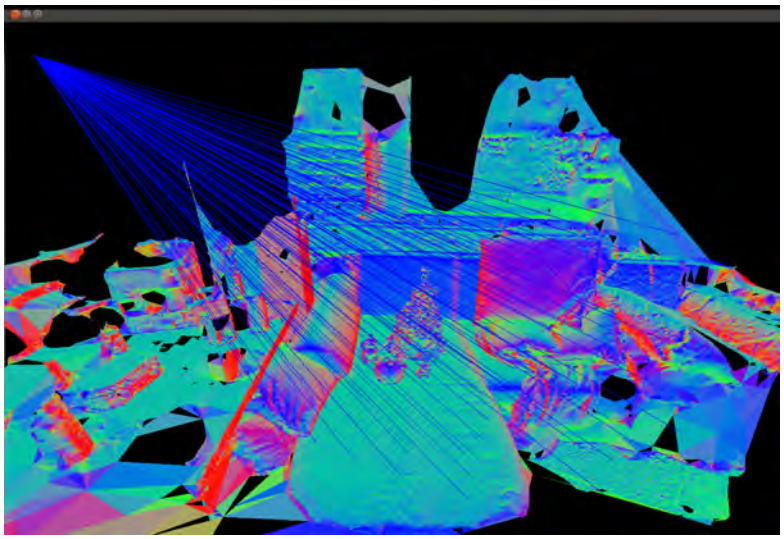
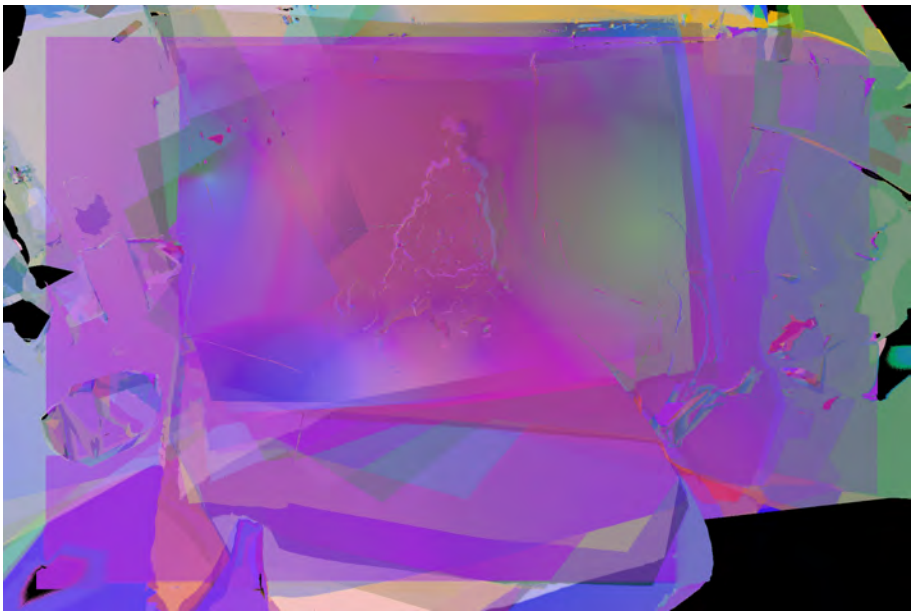


Figure 5.14: Rendering of some rays in the scene.

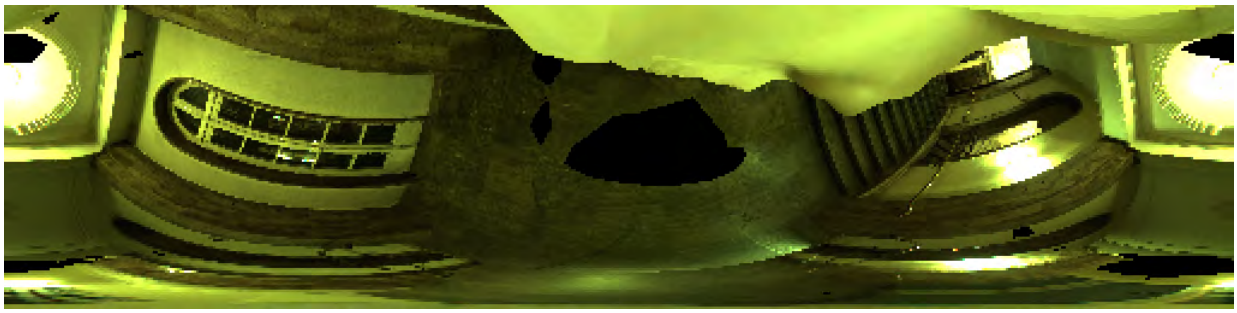
Figure 5.7. These rays begin at the same point and were shot in the scene in order to simulate a photograph of the scene under a given perspective. Some of these rays are rendered in Figure 5.14. Figure 5.15 shows the output image containing the color of each ray and its blending map, that shows how the cameras contributed for each ray. In this map, each camera is assigned to a random color and blended using the IBR method weights. It is important to notice some color discontinuities in the output image; they are due to strong differences in the exposure time of the images in the input set. Also, the position of some objects were moved during the scene acquisition, so they appear as translucent objects in the image (*e.g.* the pig on the sofa and the bottle on the table).

(a)

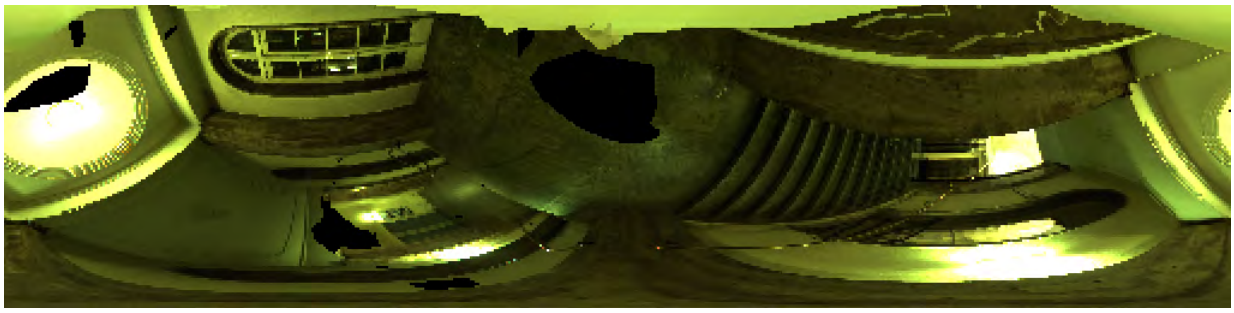


(b)

Figure 5.15: Output results from the method: (a) output image containing the color of each ray in each pixel position; (b) blending map.



(a)



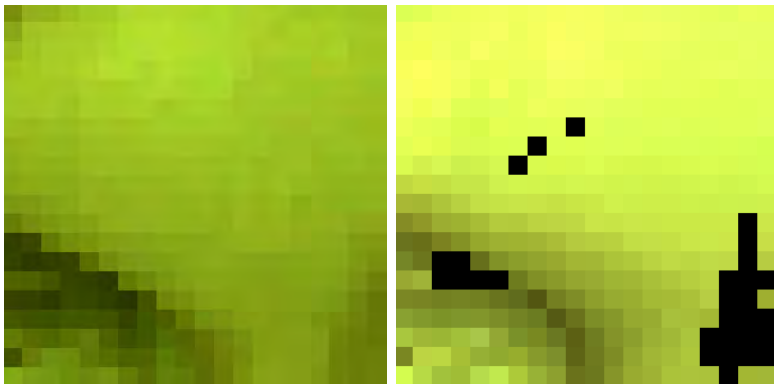
(b)

Figure 5.16: Environment maps colors, obtained: (a) from the lower left point to the hit points shown in Figure 5.10(a); (b) from the upper right point to the hit points shown in Figure 5.10(b).

The rays defined in Section 5.4 are used as input to this stage. Figure 5.16 shows two environment color maps obtained from the position of surface patch points and the maps that define where each ray should end to simulate incoming light. Notice that part of the statue is seen from the point (see the top of the figures), and that there are some holes, painted as black areas, caused by lack of geometry information. Figure 5.17 shows the results obtained during the estimation of the patch colors as seen by two virtual cameras. Points in the surface patch that are not seen by the virtual cameras are painted in black.

5.5.1 Penalties

Penalty values are used to define the color contribution of different cameras in the estimated ray color. Following the Unstructured Lumigraph Rendering approach [12], three main characteristics must be considered when evaluating the penalty values for each camera: the angular difference ($penalty_{ang}$), the resolution ($penalty_{res}$) and the field of view ($penalty_{fov}$). The final penalty value for a candidate camera i is calculated according to



(a) (b)

Figure 5.17: Viewed colors of the surface patch shown in Figure 5.8, as seen from two virtual camera positions.

Equation 5.5. The constants α, β and γ control the relative importance of each penalty, and can be set to zero if the penalty should be ignored.

$$penalty(i) = \alpha\ penalty_{ang}(i) + \beta\ penalty_{res}(i) + \gamma\ penalty_{fov}(i) \tag{5.5}$$

The angular difference is the angle between the direction of the ray and the direction from the camera i to the end of the ray. The value of the angle in radians is stored in $penalty_{ang}(i)$. However, if this angle is bigger than 90 degrees, $penalty_{ang}(i)$ receives the maximum float value to avoid that cameras that see the point in a too oblique direction contribute to the final color. The resolution penalty consider the distances from the ray start and from camera i to the end of the ray, as described in Equation 5.6 (Buehler *et al.* [12]).

$$penalty_{res}(i) = max(0, \|r_{end} - C_i\| - \|r_{end} - r_{start}\|) \tag{5.6}$$

Where r_{end} is the end point of the ray, C_i is the position of the camera that is being evaluated and r_{start} is the position where the ray starts. This penalty avoid cameras that considerably undersample r_{end}.

The field of view penalty analyses if the point is visible by a candidate camera and evaluates how close it is to its image borders. Listed below are the evaluation criteria proposed in this work for this penalty and the values each criteria add to $penalty_{fov}(i)$:

- Check if the point projection is located inside the image by projecting the point in the image using the camera i calibration matrix: if false, add the maximum possible float value and end $penalty_{fov}(i)$ estimation;
- Compare the 3D point projection with the depth map of the image from camera i and check if the point is occluded by the scene geometry: if true, add the maximum possible float value and end $penalty_{fov}(i)$ estimation;
- Check if the point is located in an invalid zone, defined by masks (*e.g.* black areas added during rectification): if true, add the maximum possible float value and end $penalty_{fov}(i)$ estimation;
- Check if the angle of view is bigger than the camera i field of view: if true, add the maximum possible float value and end $penalty_{fov}(i)$ estimation;
- Check if the point is in the boundaries of the image (defined by a threshold distance from the center of the image): increase $penalty_{fov}(i)$ based on its distance to the farthest boundary.

As discussed in Section 5.5, the color at the end of each ray gives the amount of light reflected towards the candidate cameras, and not towards the point at the start of the ray. To compensate this fact, the angular difference penalty should be higher than the others, so cameras that have similar perspective have higher contributions. This way, this method empirically assumes that α value is 0.6, while β and γ values are 0.2 each. Furthermore, only four candidate cameras are used ($k = 4$) in order to minimize the contribution of cameras with higher penalties.

5.6 SVBRDF Estimation

As a spatially variant surface is assumed, the BRDF estimation is performed per pixel (SVBRDF). Assuming that the environment map of each point and a set of observations of this point from different positions are available from the previous steps, it is possible to estimate the BRDF using several different methods. However, it is known that the data

retrieved from the previous steps is not accurate, as it is interpolated from real ones. It means that if an approach like the one made in the experiment presented in Section 4.2 were used, the resulting BRDF could be very unstable as it depends too much on the input data.

This way, an approach based on a combination of basis BRDFs (Lensch [41]) was chosen in this step. As this approach defines the BRDF as a linear combination of existing BRDFs, the results tend to be more stable and present consistent renderings. This way, a BRDF can be defined as the sum of a Lambertian component with albedo $\frac{c}{\pi}$ and $N-1$ specular terms $f_k(\omega_i, \omega_o)$. The Lambertian component represents the diffuse contribution, which is the light that is equally reflected at all directions. The directional reflectance is modeled by the weighted sum of the N-1 specular terms (see Equation 5.7). Assuming that these specular terms are known, the problem is reduced to find the N values for a_n $(0 \leq n < N)$. As colors are considered only additive, the values of a_n must never be smaller than zero and can be found through a non-negative linear least-squares optimization.

$$f_r(\omega_i, \omega_o) = a_0 \frac{c}{\pi} + \sum_{k=1}^{N-1} a_k \ f_k(\omega_i, \omega_o). \tag{5.7}$$

In this work, an isotropic version of the Ashikhmin-Shirley model [7] is used to represent the specular component, using the derivation presented by Ngan *et al.* [55], described in Equation 5.8. The reasons for using this model instead of the Lafortune model (adopted in the experiment described in Section 4.2) are that the Lafortune model is not physically-based (refer to Subsection 2.3.3.1) and it presented convergence problems during the experiments performed in this thesis (Subsection 4.2.3). Also, the Ashikhmin-Shirley model has been evaluated by Ngan *et al.* [54], presenting one of the best results amongst the evaluated BRDF models, whose included the Lafortune model (see Section 3.5).

$$f_k(\omega_i, \omega_o) = \frac{m_k + 1}{8\pi} \frac{(n \cdot h)^{m_k}}{(\omega_o \cdot h) max((n \cdot \omega_i), (n \cdot \omega_o))} F_k(\omega_o \cdot h). \tag{5.8}$$

Where h is the half angle vector between ω_i and ω_o and m_k is a Phong-based exponent that control the shape of the specular lobe for isotropic surfaces. The Fresnel factor F_k is defined using the Schlick's approximation [62] (Equation 5.9), where R_k represents the reflectance of the material at normal incidence and θ is the angle between the viewing direction and the half-angle direction $(\omega_o \cdot h)$. This way, it is correct to assume that the shape of the specular lobe is parameterized by m_k and R_k.

$$F_k(\theta) = R_k + (1 - R_k)(1 - cos(\theta))^5. \tag{5.9}$$

The first step in the BRDF estimation is to define which are the parameters of the Ashikhmin-Shirley model that will define the basis BRDFs. Once they are defined, they are used to render the specular terms of the linear combination of the BRDFs, and this linear combination is solved in order to find its coefficients $(a_0, a_1, ..., a_n)$. These variables define the BRDF basis linear combination that can be used to render and evaluate the appearance of each point. The remaining of this section describes the main aspects of this approach.

5.6.1 Basis BRDFs Selection

An approach similar to the one proposed by Häußler [31] is adopted to select the basis BRDFs (refer to Section 3.2) for each point in the patch. This approach consists in clustering the MERL Ashikhmin-Shirley BRDFs [51, 55] in a 2D space defined by their m and R values. The MERL data set is composed of 100 BRDFs of different materials, including metals, plastics, painted surfaces and fabrics. This data set does not cover the full range of materials. However, linear combinations of representative functions in this data set generate a spam of possibilities that overcome the original number of BRDFs and have the potential to represent a broader range of materials.

The cluster method used is the *k-means*, and only lobes with $m > 5$ are used, as materials with smaller m are considered fully diffuse and can be approximated by the diffuse component of the BRDF function. As the m values present a high variation in

their values, the author scales them down by using their cubic root during the clustering. The centroid of each resulting cluster is used as a basis BRDF.

The main motivation in the clustering is that if all MERL BRDFs were used as basis, the BRDF estimation and rendering would result in longer computation times and higher memory requirements. So, a set of representative values is extracted using the clustering approach and used as basis BRDFs. The approach adopted here is slightly different, as it normalizes both m and R in $[0, 1]$ instead of taking the cubic root of m. This way, a proportional transformation is done in both values.

The number of clusters will define the number of basis BRDFs to be used in the following steps of the proposed method. Tests were performed by fitting the MERL BRDFs as a combination of basis BRDFs retrieved from different numbers of clusters (5, 10, 20, 30, 40, 50, 60 and 70). The evaluation is done using the same approach adopted by Ngan *et al.* [54] and Häußler [31]. This approach minimizes the difference between the original BRDF values and the ones estimated using the linear combination of the clustered coefficients and by normalizing the resulting values for each channel by the maximum albedo of the BRDF. This normalization enables the comparison of the residual errors for different materials while discounting the differences in their relative scales.

Figure 5.18 shows a plot with the mean normalized errors for the color channels at each cluster configuration. It compares the errors obtained using the approach adopted by Häußler [31] (m values are scaled down by their cubic root) with the approach presented in this work (m and R are normalized between 0 and 1). The error difference between these two approaches is higher when less than 20 clusters are evaluated, getting similar when the number of clusters increases.

Based on the evaluation performed, the approach proposed in this work with a number of 20 clusters presents a good tradeoff between the mean error and number of basis BRDFs. This configuration is used to represent the basis BRDFs for synthetic scenes in Section 5.7. However, the higher the number of degrees of freedom in the model, the smaller the fitting error would appear in the experiments with real scenes. Also, whenever the number of degrees of freedom in the model is higher than or equal to the number of samples (real

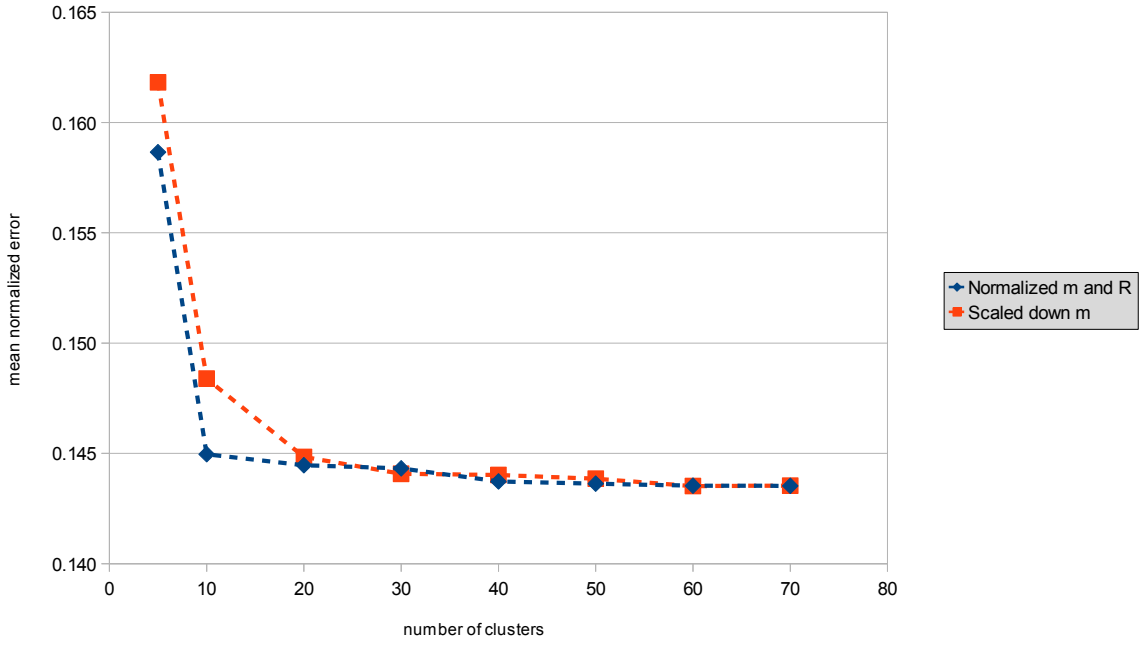


Figure 5.18: Mean error for a different number of clusters using: (square series) the approach presented by Häußler [31]; (diamond series) the approach presented in this work.

images), perfect fit with zero error is possible. However, the risk of fitting noise in the captured data (*i.e.* overfitting) is also high. In this case one would also model random fluctuations in the camera sensor, instead of only the actual BRDF of the scene.

These facts raise need to adjust the tradeoff between fitting error and model complexity. To this end, leave-one-out cross-validation is performed in order to find an optimal number of BRDF parameters for real scenes. In this approach, one of the input views is kept aside, for testing, while the other views are used to fit the BRDF model with 5, 10, 15, 20 and 30 parameters (clusters). The second step consist in analyzing how well the learned model can predict the object appearance in the view that was left aside. To this end, Normalized Root Mean Squared Error (NRMSE) is calculated for each pair of left-aside input view and its predicted values. This procedure is repeated for all input views and the prediction errors are averaged and used to select the adequate number of parameters for each patch.

5.6.2 Assemble and Solve Linear Combinations

To solve the linear combination defined in Equation 5.7 and retrieve the values of the coefficients a_n it is necessary to calculate samples for c and $f_k(\omega_i, \omega_o)$ for each point to be evaluated. As the BRDF is calculated per point, the c value is calculated once per point, and $f_k(\omega_i, \omega_o)$ is calculated based on each basis BRDFs and sampled directions. The calculation of these components is described in Subsection 5.6.3.

With these values it is possible to assemble one matrix containing the observed color samples for each point and another containing the corresponding values of c and $f_k(\omega_i, \omega_o)$ for each sample. This way, assuming that the number of samples is equal or bigger than N, is is possible to retrieve the a_n coefficients using a linear squares optimization. As the values of a_n must never be smaller than zero, the non-negative linear least-squares optimization method developed by Lawson and Hanson [40] is used. The optimization is performed at each color channel of the samples, so there will be a resulting set of BRDF coefficients for each channel.

5.6.3 BRDF Components

As stated in Section 5.6, the first coefficient (a_0) defines how much the diffuse component c will contribute to the point color. Based on the definition of irradiance at a point (Equation 2.6), c is calculated as the integration of the incoming light from all positions of the environment map L_i, weighted by the cosine of the angle between the the normal vector of the point and the incident light direction (θ_i), as formalized in Equation 5.10.

$$c = \int_{\mathcal{H}^2(\mathbf{n})} L_i(\mathrm{p}, \omega_i) cos\theta_i \mathrm{d}\omega_i. \tag{5.10}$$

The a_k remaining coefficients of Equation 5.7 weights the specular components of the BRDF ($f_k(\omega_i, \omega_o)$). As the specular reflection depends on the direction of the camera, some incident light directions present more influence than others in the final BRDF value retrieved by $f_k(\omega_i, \omega_o)$. This influence can be measured through the Ashikhmin-Shirley Probability Density Function (PDF) (Ashikhmin and Shirley [7]), derived from

the Ashikhmin-Shirley model shown in Equation 5.8. This way, given a view direction and a random input light direction, this function will estimate the weight $w(\omega_i, \omega_o)$ of this specific incoming light direction in the final BRDF value (Equation 5.11).

$$w(\omega_i, \omega_o) = \frac{1}{max((n \cdot \omega_i), (n \cdot \omega_o))} F(F_0, \omega_o \cdot h)(n \cdot \omega_i). \tag{5.11}$$

Monte Carlo integration with importance sampling is used, following the approach developed by Häußler [31]. The hemisphere around the normal vector of the point is sampled using the Ashikhmin-Shirley probability density function and integrated based on the weights in order to retrieve the BRDF value for $f_k(\omega_i, \omega_o)$. This ensures fast convergence with a smaller number of samples.

5.7 Results and Evaluation

Surface patches were sampled and their SVBRDFs estimated using the method described in this work. To evaluate the quality of the results, two approaches were used. In the first one, reference views of a patch are generated from virtual camera perspectives. The proposed SVBRDF estimation method is applied on this patch and linear combinations of BRDF models are found for each point in the patch. Images of the patch are then rendered from the same perspective of the virtual cameras using the BRDFs found and compared with the views of the patch.

In the second approach, a random combination of BRDFs is generated for each point in a surface. Synthetic reference views of this patch are generated based on these combinations and used as input to the method. The estimated BRDF combinations retrieved by the method are used to render images of the patch from the same perspective as the synthetic reference views. These two views (reference and estimated) are then compared in order to measure the performance of the method in ideally sampled situations, as the synthetic reference views will always contain ideal data.

The rendered images are obtained using Monte Carlo integration to render colors for each point from its environment map and its estimated BRDF coefficients. The first

approach shows how reliable the method is when views retrieved using unstructured lumigraphs are used. The second approach measures how precise the method is when views of a patch that are coherent with each other (*i.e.* have the same basis BRDF combination at each point) are used. The second approach shows a perfect case for the proposed method; when enough color information of each point is obtained through the input HDR images, and the right color variation is stored in each view using the light field interpolation approach.

An error HDR image is generated for each comparison between estimated and rendered images. It contains the absolute difference between the values stored for each point in the surface patch. The error images were tone mapped to be shown in this work. This way, they quantify when the error in one point is higher than the error in another point in the same image; the color intensities cannot be used to compare points in different images. To evaluate this difference, the Normalized Root Mean Squared Error (NRMSE) is calculated for each pair of reference and estimated views of a patch. They are normalized by the difference between the maximum and minimum intensities found for each channel in all the reference images. The average and pattern deviation of the NRMS errors of all views are shown for each patch, as well as their minimum and maximum values.

It is important to mention that the number of virtual cameras is chosen in order to avoid over and under-fitting of the BRDF model. Tests were performed in order to evaluate the performance of the method using only data from the real cameras, without the camera interpolation step performed through unstructured lumigraph rendering. The errors were consistently higher when real cameras were used. Although these cameras contain potentially more precise information, their sample is not as regular and the random fluctuations in the camera sensor affect the behavior of the estimated model by inserting noise in its predictions. Certainly more densely sampled data would generate better models from the real observations. Nevertheless, the purpose of this approach is to be able cope with sparse data, and to provide a more flexible and intuitive way of looking at the problem.

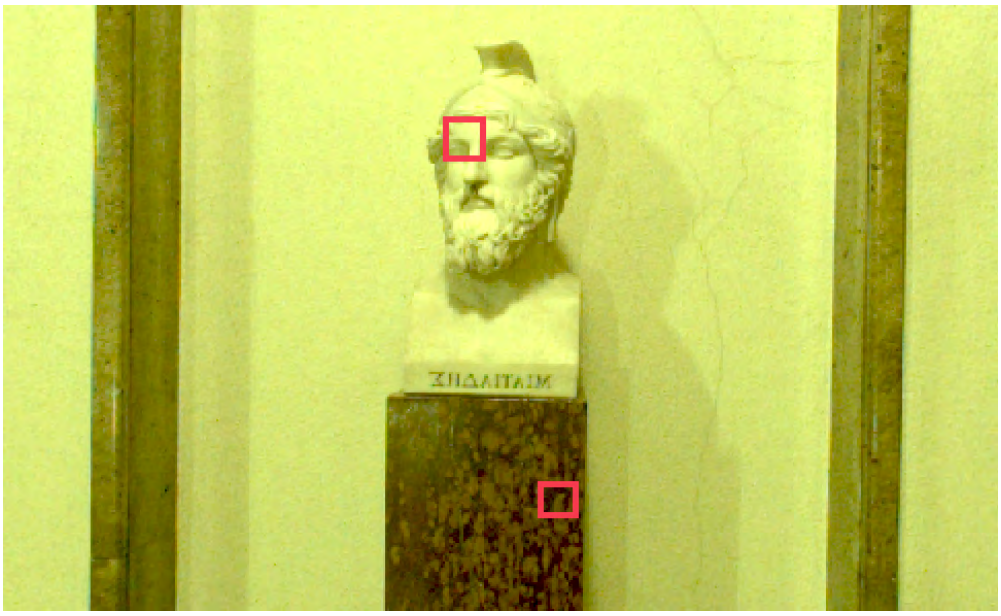


Figure 5.19: Location of the patches in the Neue Aula scene.

This way, only virtual cameras were used to estimate and evaluate the BRDF of patches in the scenes described in Subsection 5.2.1. Subsection 5.7.1 shows results obtained for the *neue_aula* dataset, and Subsection 5.7.2 shows results for the *office* dataset.

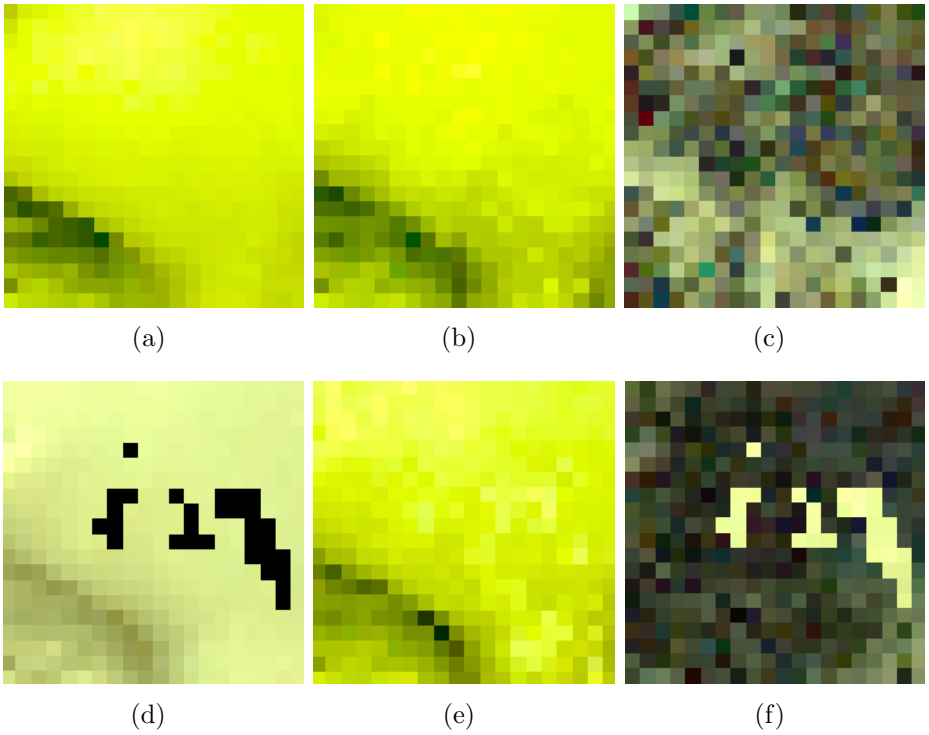
5.7.1 Neue Aula Scene

The Neue Aula scene presents many details and constant lighting, which provided a good geometry reconstruction and hereby improved the proposed method results. This scene is composed of 22 sets of six images from the Ladybug camera, resulting in 132 HDR images of the scene. Each image has 2048×2048 pixels.

Results for two real patches and one synthetic patch are shown. The first real patch is located in the forehead of a statue in the scene, and the second one is located in its basis, as highlighted in Figure 5.19. The synthetic patch uses geometry and lighting from the patch at the statue's basis. The number of virtual cameras and environment maps used for each patch is presented in Table 5.1, together with the resolution (number of sampled light rays) of the patch, the resolution of each environment map and number of basis BRDFs used. The number of BRDFs was chosen based on the methodology described in Section 5.6.1.

Table 5.1: Data used in the Neue Aula scene patches.

	Forehead	**Basis**	**Synthetic**
Patch resolution	400 (20×20)	64 (8×8)	64 (8×8)
Environment map resolution	65536 (512×128)	65536 (512×128)	65536 (512×128)
Number of virtual cameras	36	23	23
Number of environment maps	400	64	64
Number of basis BRDFs	30	10	20



(a) (b) (c)

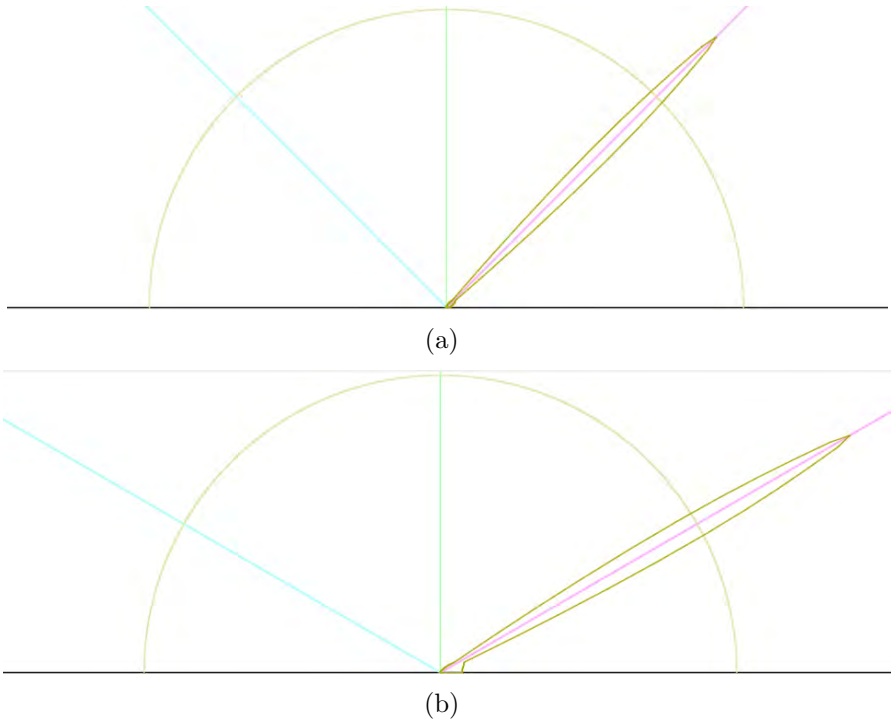
(d) (e) (f)

Figure 5.20: Results obtained for two views of the patch located in the statue's forehead: (left) views from virtual cameras; (center) rendering of patch from same view; (right) image containing the difference between the first two columns.

Figure 5.20 shows results for two views of the patch located in the statue's forehead. Information about the NRMS error of all the virtual cameras used in this patch are shown in Table 5.2. It is relevant to notice that Figure 5.20(d) depicts a partial view of the patch, where the black points were not seen by the virtual camera. In the rendered image (Figure 5.20(e)) these points are rendered in a consistent way, showing that the method generate plausible results when values that were not given as input are reproduced. The difference values in these positions are not used in the calculation of the error values. Figure 5.21 shows polar plots of the material estimated for the lower right point in the patch (see Figure 5.20) under two different incident angles.

Table 5.2: Normalized Root Mean Squared Error calculated for *Forehead* patch.

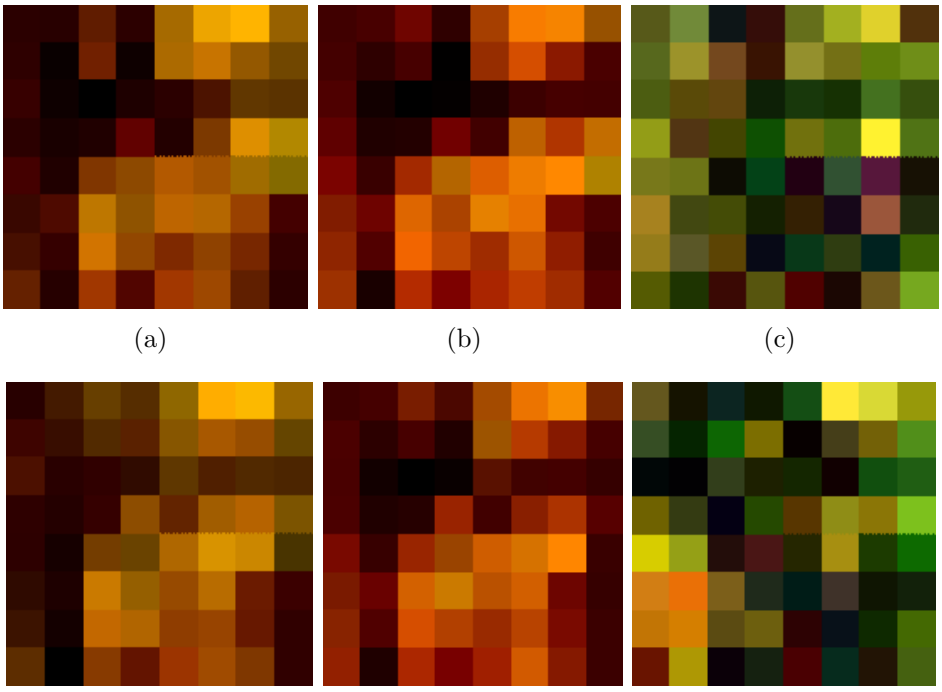
	R	**G**	**B**
Average NRMSE	0.0666	0.0694	0.0639
Minimum NRMSE	0.0019	0.0021	0.0013
Maximum NRMSE	0.1453	0.1480	0.1245
NRMSE pattern deviation	0.0354	0.0367	0.0329



(a)

(b)

Figure 5.21: Polar plot of the luminance generated by the estimated BRDF under different incident angles (θ_i): (a) 45 degrees; and (b) 60 degrees. The plots were generated using Disney's BRDF Explorer tool [20].



(a) (b) (c)

(d) (e) (f)

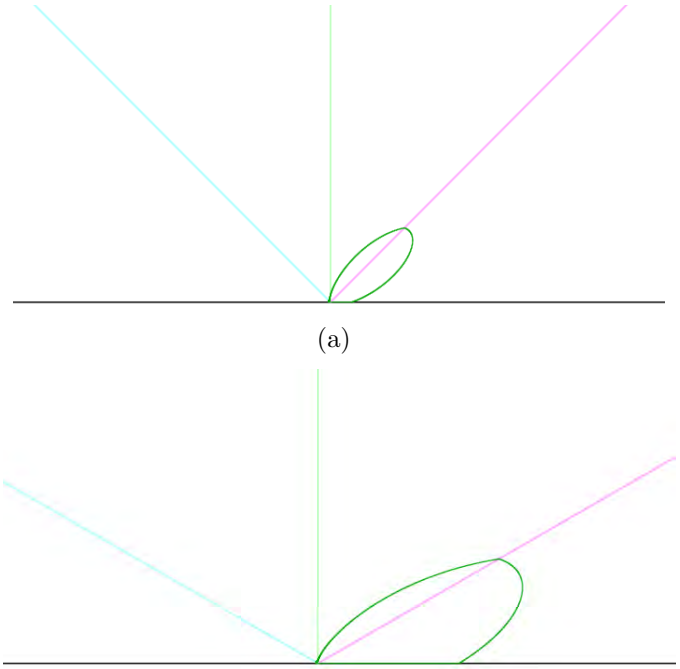
Figure 5.22: Results obtained for two views of the patch located in the basis of the statue: (left) views from virtual cameras; (center) rendering of patch from same view; (right) image containing the difference between the first two columns.

Table 5.3: Normalized Root Mean Squared Error calculated for *Basis* patch.

	R	**G**	**B**
Average NRMSE	0.1303	0.1401	0.1256
Minimum NRMSE	0.0721	0.0707	0.0729
Maximum NRMSE	0.2334	0.2434	0.1558
NRMSE pattern deviation	0.0401	0.0421	0.0234

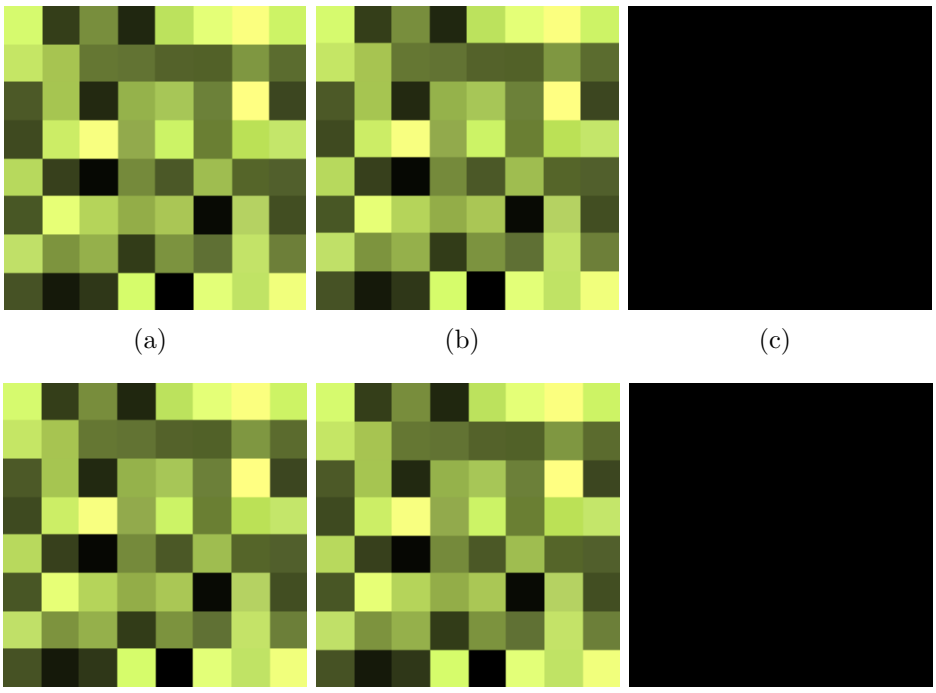
Figure 5.22 shows the results for the first two views of the patch located in the basis of the statue. Information about the NRMS error of all the virtual cameras used in this patch are shown in Table 5.3. Figure 5.23 shows polar plots of the material estimated for the lower right point in the patch (see Figure 5.22) under two different incident angles.

Figure 5.24 shows the results for the first two synthetic views of a patch made using the same geometry and incident lighting as the patch in Figure 5.22. Information about the NRMS error of all the virtual cameras used in this patch are shown in Table 5.4.



(a) (b)

Figure 5.23: Polar plot of the luminance generated by the estimated BRDF under different incident angles (θ_i): (a) 45 degrees; and (b) 60 degrees. The plots were generated using Disney's BRDF Explorer tool [20].



(a) (b) (c)

(d) (e) (f)

Figure 5.24: Results obtained for two views of the same patch: (left) synthetic input images; (center) rendering of patch from same view; (right) image containing the difference between the first two columns.

Table 5.4: Normalized Root Mean Squared Error calculated for *Synthetic* patch in Neue Aula scene.

	R	**G**	**B**
Average NRMSE	1.45344e-07	1.51955e-07	1.17898e-07
Minimum NRMSE	1.05000e-07	1.11383e-07	9.02269e-08
Maximum NRMSE	3.67527e-07	3.98801e-07	2.91424e-07
NRMSE pattern deviation	5.27780e-08	5.81325e-08	4.08337e-08

Table 5.5: Data used in the Office scene patches.

	Speaker	**Synthetic**
Patch resolution	100 (10×10)	100 (10×10)
Environment map resolution	65536 (512×128)	65536 (512×128)
Number of virtual cameras	32	32
Number of environment maps	100	100
Number of basis BRDFs	10	20

5.7.2 Office Scene

The office data set was taken in a room with white walls and furniture and blue floor (see Figure 5.4). Big areas with constant colors are a problem in the geometry reconstruction stage, as a reduced number of features decreases the quality of the geometry reconstructed using the SFM+MVS approach. This way, this scene's geometry presents a considerable amount of noise and missing areas. This fact caused the reduction of the areas where patches could be selected (the patch must be located in an area with smooth and more precise geometry).

This scene is composed of 30 sets of six images from the Ladybug camera, resulting in 180 HDR images of the scene. Each image has 1024×1024 pixels. The patch evaluated in this subsection is located on a speaker, highlighted in Figure 5.25. A synthetic patch uses geometry and lighting from this patch. The number of virtual cameras and environment maps used for each patch is presented in Table 5.5, together with the resolution of the patch, the resolution of each environment map and number of basis BRDFs used. The number of BRDFs was chosen based on the methodology described in Section 5.6.1.

Figure 5.26 shows results for two views of a patch located in the speaker area. Information about the NRMS error of all the virtual cameras used in this patch are shown in Table 5.6. Once again, it is important to notice that Figure 5.26(a) depicts a partial view

Figure 5.25: Location of the patch in the office scene.

Table 5.6: Normalized Root Mean Squared Error calculated for *Speaker* patch.

	R	**G**	**B**
Average NRMSE	0.1348	0.1246	0.1082
Minimum NRMSE	0.0469	0.0438	0.0354
Maximum NRMSE	0.2860	0.2486	0.1914
NRMSE pattern deviation	0.0567	0.0484	0.0377

of the patch, and that the missing points colors were inferred in a consistent way in the rendered image (Figure 5.26(b)). Figure 5.27 shows polar plots of the material estimated for the lower right point in the patch (see Figure 5.26) under two different incident angles.

Figure 5.28 shows the results of two synthetic views of a patch made using the same geometry and incident lighting as the patch in Figure 5.26. Information about the NRMS error of all the virtual cameras used in this patch are shown in Table 5.7.

Table 5.7: Normalized Root Mean Squared Error calculated for *Synthetic* patch in Office scene.

	R	**G**	**B**
Average NRMSE	1.17673e-07	1.29739e-07	1.22784e-07
Minimum NRMSE	8.49215e-08	9.84799e-08	9.52174e-08
Maximum NRMSE	3.40649e-07	3.63463e-07	3.45201e-07
NRMSE pattern deviation	4.37528e-08	4.55104e-08	4.31577e-08

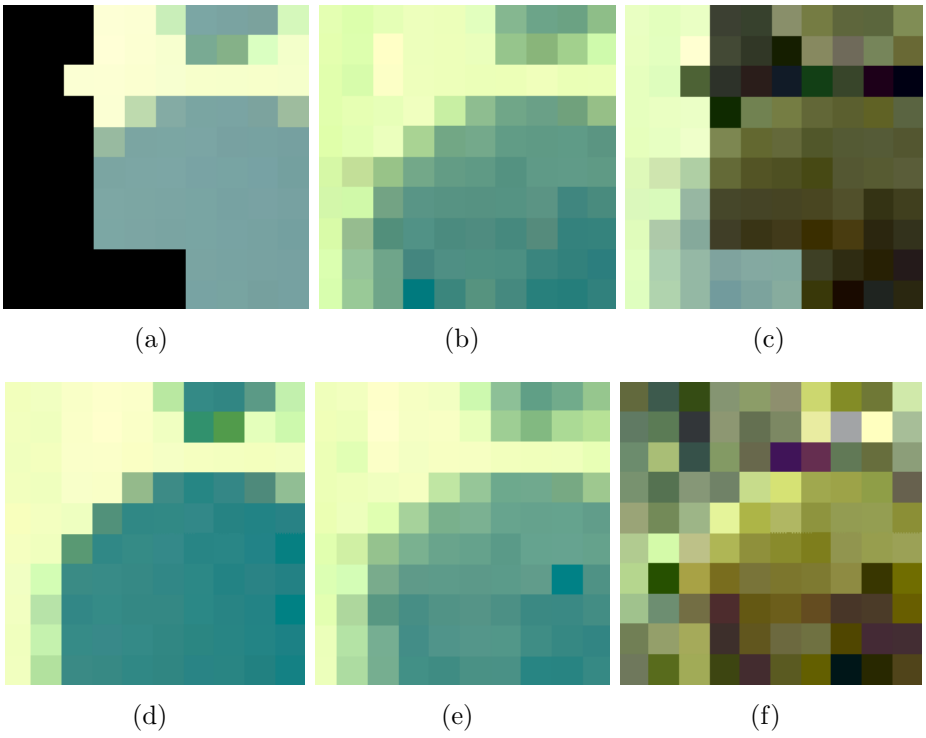


Figure 5.26: Results obtained for two views of the patch located on the speaker: (left) views from virtual cameras; (center) rendering of patch from same view; (right) image containing the difference between the first two columns.

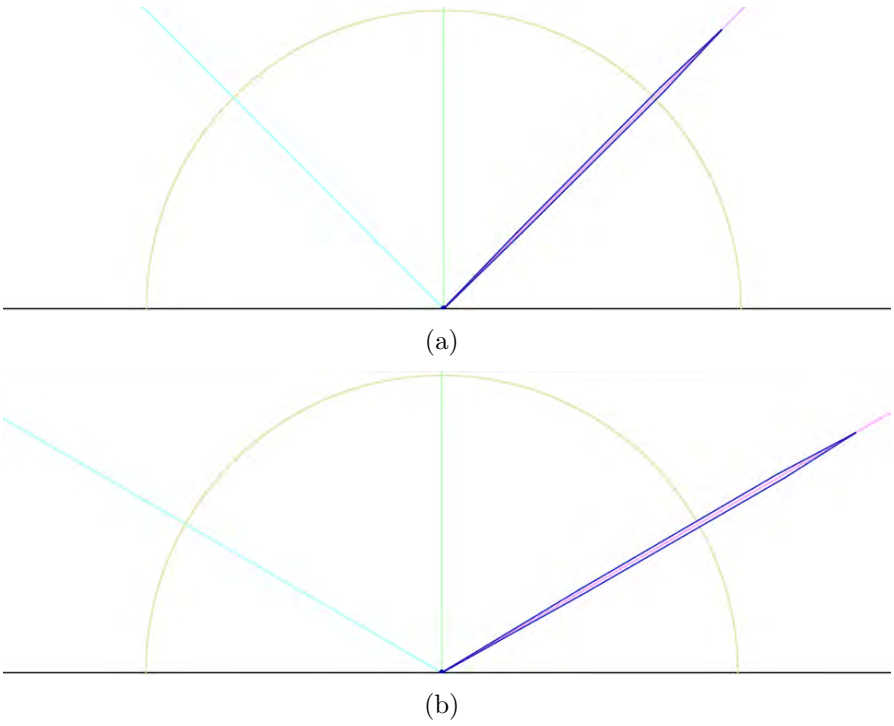


Figure 5.27: Polar plot of the luminance generated by the estimated BRDF under different incident angles (θ_i): (a) 45 degrees; and (b) 60 degrees. The plots were generated using Disney's BRDF Explorer tool [20].

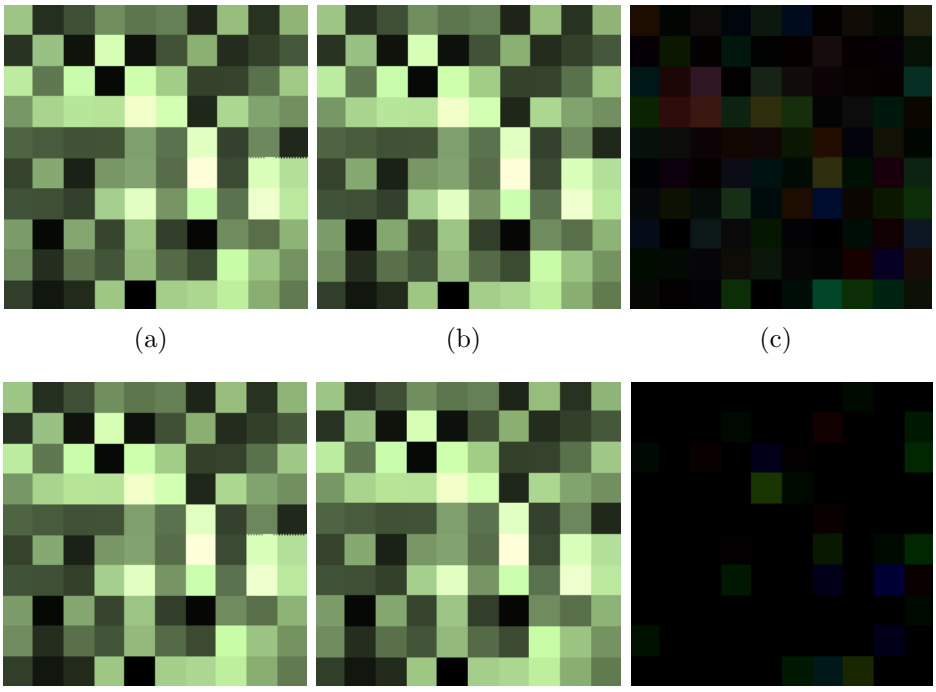


Figure 5.28: Results obtained for two views of the patch located on the speaker: (left) synthetic input images; (center) rendering of patch from same view; (right) image containing the difference between the first two columns.

5.8 Conclusions

The proposed method was designed to consider the desirable properties for an image-based preservation method delineated in Section 3.6. These properties have been proposed based on the literature review performed in Chapter 3, and describe key aspects that simplify the acquisition setup, reduce the storage requirements and model the appearance with realism.

To this end, the proposed image-based preservation method estimates the SVBRDF of a surface patch in a scene based on omnidirectional HDR images of this scene, simplifying the acquisition step. The geometry of the scene is reconstructed from the HDR images, using a combination of Structure from Motion and Multi View Stereo methods. This geometry is then used as reference to estimate rays with desired behavior, in a way that these rays colors correspond to samples of the environment map and observations of the points in the patch. These samples are then used in the BRDF estimation of each point in the surface patch, providing a spatially variant BRDF estimation.

The BRDF of each point is represented as a linear combination of a diffuse term and specular terms, where the specular terms are represented using Ashikhmin-Shirley models. The use of basis BRDFs ensure a compact representation, as the only stored parameters are the coefficients of the linear combination of basis BRDFs for each point and the basis BRDFs. The use of Ashikhmin-Shirley models as specular terms increases the reliability of the BRDF representation, as this model was in the group that obtained the best results in the BRDF comparison performed by Ngan *et al.* [54] (refer to Section 3.5).

The leave-one-out cross-validation showed that the statue basis and the office speaker patches presented the smallest prediction errors with 10 basis BRDFs. On the other side, the statue forehead patch presented the best results with 30 basis BRDFs, indicating that this test is important to analyze the properties of the material. It is important to mention that the optimized number of basis BRDFs led to smaller errors during the BRDF estimation with all views.

The results obtained in each step of the method are shown in this chapter together with an evaluation of the resulting SVBRDFs. The results are very satisfactory, especially when the extreme interpolated nature of this method is considered. It is interesting to notice that the quality of the results increases with the number of images, so if the user is interested in some specific part of the scene, he/she can take more pictures of this part, providing more information about this area. This extra information will improve all the stages of the method, from the geometry reconstruction to the SVBRDF estimation. The main demand of the method regarding the number of images is that a regular sample of the scenario must be imaged, so that the images can be used to estimate the geometry and light rays colors.

The main purpose of this work is to propose an original approach that is flexible and retrieves consistent results. This way, its next research directions would be to reduce computing costs (Subsection 5.8.1) and to cope with the limitations discussed in Subsection 5.8.2. Assuming perfect data and dielectric materials, only one scene could be enough to capture the BRDFs of a surface patch, as the illumination would be known and its color and intensity would be subtracted from the scene. As the data is not perfect

and the physical formulation of the problem is an approximation of a phenomenon that happens at a microscopic level, the resulting BRDFs can always present some influence of the environment lighting. This influence could be reduced by the acquisition of the same scene under different environments.

Another possible direction is to improve the SVBRDF estimation approach by clustering the points in the scene according to the material they are composed of. This would reduce the noise during rendering and the storage requirements. Also, a quantitative comparison against a similar approach that uses a light probe (chrome sphere) or a laser scanner could be performed, in order to capture the environment map and the scene geometry more accurately. This data can be used to isolate and analyze the errors introduced by the different stages and identify more opportunities for improvement. This experiment could also be simulated via the rendering of a synthetic scene.

5.8.1 Computing Costs

Several decisions influence the overall computing cost of the proposed method. The geometry reconstruction is influenced by the number of acquired images, as the more images are acquired, the more time and processing are spent in this stage. This figures as a valid cost since the geometry is reconstructed only once and more images enable the generation of more precise 3D models.

The ray tracing stage processing time is defined by the resolution of the patch and of the environment maps, as one environment map is generated for each point in the patch. The resolution and number of those images is crucial during the light ray color estimation, as the unstructured lumigraph approach needs to interpolate several images in order to retrieve the color of each ray. In this stage the number of input images and the value of k (candidate images) increase the storage and processing requirements, as penalty values are generated for all the input cameras before the selection of the k best ones.

As Monte Carlo integration is used during the rendering of the specular BRDF components, the processing time depends on the number of samples that must be used to integrate the final value. As the PDF of the Ashikhmin Shirley model is used, a reduced

number of samples can be used to retrieve converging and fast results (*e.g.* 1000 samples were used in the experiments presented in this chapter). In this scenario, the main cost in the BRDF representation comprises the memory requirements, as more basis BRDFs increases the required storage space.

5.8.2 Limitations

Although processing is a current relevant cost, the method can be easily parallelized as the calculations are performed in a per pixel basis. Also, there is a vast field for optimizations using the Graphics Processing Unit and multi threading. To cope with this limitation, only small patches were used in the experiments performed in this chapter.

The tone mapping from HDR to LDR images required by the MVS stage during geometry reconstruction causes a heavy lost of details in the HDR image and thus affects the resulting geometry by reducing its resolution and quality. The utilization of a MVS method that accepts HDR images as input would improve the geometry considerably and thus the results obtained in this work.

Holes in the geometry are represented as black pixels in the environment maps, and this error is spread into the SVBRDF. Although the holes in the scenes shown in this chapter are located in places that are not crucial to infer the incoming light, this can be a problem if the hole is located in an area where there is an important light source. Hole filling algorithms could be a solution to this problem, and as the obtained geometry can be composed of different parts and contain noise, this would be a challenging case study. As it is assumed that the patch does not contain holes, this problem does not apply to the rays that simulate virtual cameras.

CHAPTER 6

CONCLUSIONS

This thesis presented a research regarding the use of images in the preservation of the appearance of real world objects. It begins by introducing fundamental concepts about basic elements on the comprehension of the behavior of light and its representation. After that, a literature review compiles works that aim to capture and represent the appearance of real objects. This review focuses on image-based methods, so different types of images are considered. These works are compared and discussed, in order to recognize and categorize their main characteristics. Based on this study, two main experiments were performed in order to understand and apply appearance preservation concepts. The first experiment reduces at most the influence of lighting and the second one assumes directional lighting, using a single light source. The research made and the experiments performed established grounds to the development of a novel method.

Image-based appearance preservation methods are very demanding, both in data acquisition efforts and in the required processing pipeline. A great number of images from different positions must be taken to capture the appearance of data, and when dealing with HDR images this number increases even more. The processing is very challenging, as the final appearance of the object is influenced by the sequence of steps chosen by the method, steps which also perform non trivial and error prone tasks such as geometry reconstruction.

Three main approaches have been discussed in this thesis regarding how light is considered in image-based appearance preservation methods. As shown in the first two groups of experiments performed in this thesis, to disregard light or to consider only direct lighting are approaches that demands highly controlled environments and acquisition setups. Moreover, these approaches generate good but limited results because they ignore all or

part of the incident lighting, which may contain data that are relevant to the definition of the object appearance.

Based on the results obtained with the previous experiments, a novel method in the image-based appearance preservation is proposed. This method considers the whole environment as source of light, which is an approach that enables a more flexible acquisition setup and assumes more complete illumination information about each point. As a huge number of data is required to capture all the information involved, this approach is not as explored as the other two. In this scenario, the method proposed in this thesis presents a novel contribution. As incoming light changes from point to point, it uses unstructured lumigraph rendering to generate local environment maps for each point in the patch. These maps are used to generate space variant reflectance models for the region of interest, reproducing what happens in reality.

The quality of the results obtained from the proposed method increases with the number of images, as the problem is defined as sets of interpolated light rays. It is important to notice that, even with a relatively small number of samples, it presents consistent results. This method shows that considering the whole environment as source of light is a viable approach to obtain reliable results and to enable a more flexible acquisition setup.

This proposed method is original and it is very challenging as it opens a series of aspects to be worked and further investigated. Possible next directions are the development of techniques to improve its steps and to overcome its current limitations.

REFERENCES

[1] Neil Alldrin, Todd Zickler, and David Kriegman. Photometric stereo with non-parametric and spatially-varying reflectance. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1 –8, 2008.

[2] Beatriz Trinchão Andrade. Utilizando fotografias digitais de alta qualidade na geração de textura para modelos 3D: uma abordagem prática na preservação digital de acervos culturais e naturais. Master's thesis, Universidade Federal do Paraná, Campus Centro Politécnico, Curitiba, Paraná, February 2009.

[3] Beatriz Trinchão Andrade, Olga Regina Pereira Bellon, Luciano Silva, and Alexandre Vrubel. Enhancing color texture quality of 3D models for digital preservation of Indigenous ceramic artworks. In *Workshop on eHeritage and Digital Art Preservation. IEEE International Conference on Computer Vision*, pages 980–987, 2009.

[4] Beatriz Trinchão Andrade, Olga Regina Pereira Bellon, Luciano Silva, and Alexandre Vrubel. Digital preservation of Brazilian indigenous artworks: Generating high quality textures for 3D models. *Journal of Cultural Heritage*, 13(1):28 – 39, 2012.

[5] Beatriz Trinchão Andrade, Caroline Mazetto Mendes, Jurandir de Oliveira Santos Jr., Olga Regina Pereira Bellon, and Luciano Silva. 3D preserving XVIII century barroque masterpiece: Challenges and results on the digital preservation of Aleijadinho's sculpture of the Prophet Joel. *Journal of Cultural Heritage*, 13(2):210 – 214, 2012.

[6] Beatriz Trinchão Andrade, Caroline Mazetto Mendes, Dyego Rogher Drees, Alexandre Vrubel, Olga Regina Pereira Bellon, and Luciano Silva. Preservação digital de acervos naturais e culturais na UFPR: Iniciativas do grupo IMAGO para a construção de modelos 3D realísticos e museu virtual interativo acessível pela Internet. *Revista Extensão em Foco*, 247:87 – 99, 2010.

[7] Michael Ashikhmin and Peter Shirley. An anisotropic phong BRDF model. *Journal of Graphic Tools*, 5:25–32, Fevereiro 2000.

[8] Gary A. Atkinson and Edwin R. Hancock. Two-dimensional BRDF estimation from polarisation. *Computer Vision Image Understanding*, 111:126–141, Agosto 2008.

[9] Christian Bailer, Manuel Finckh, and Hendrik P.A. Lensch. Scale robust multi view stereo. In *Computer Vision – ECCV 2012*, Lecture Notes in Computer Science. Springer, 2012.

[10] James F. Blinn. Models of light reflection for computer synthesized pictures. *ACM SIGGRAPH Computer Graphics*, 11:192–198, Julho 1977.

[11] Rod Bogart, Florian Kainz, and Drew Hess. *The OpenEXR file format*. GPU Gems, 2004.

[12] Chris Buehler, Michael Bosse, Leonard McMillan, Steven Gortler, and Michael Cohen. Unstructured lumigraph rendering. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '01, pages 425–432, New York, NY, USA, 2001. ACM.

[13] Vlado Cerný. Thermodynamical approach to the traveling salesman problem: an efficient simulation algorithm. *Journal of Optimization Theory and Applications*, 45(1):41–51, 1985.

[14] Robert L. Cook and Kenneth E. Torrance. A reflectance model for computer graphics. *ACM Transactions on Graphics (TOG)*, 1:7–24, Janeiro 1982.

[15] Massimiliano Corsini, Matteo Dellepiane, Federico Ponchio, and Roberto Scopigno. Image-to-geometry registration: a mutual information method exploiting illumination-related geometric properties. *Computer Graphics Forum*, 28(7):1755–1764, 2009.

[16] Aldo Cumani. Edge detection in multispectral images. *Computer Vision Graphics and Image Processing: Graphical Model and Image Processing*, 53(1):40–51, 1991.

[17] Aldo Cumani. Efficient contour extraction in color images. In *Asian Conference on Computer Vision*, pages 582–589, 1998.

[18] Elodie de Mondenard. Mise en oeuvre concrète des récentes normes ISO appliquées à la photographie numérique, 2006. Mémoire de fin d'Études - Recherche Apliquée.

[19] Matteo Dellepiane, Marco Callieri, Federico Ponchio, and Roberto Scopigno. Mapping highly detailed color information on extremely dense 3D models: the case of David's restoration. *Computer Graphics Forum*, 27(8):2178–2187, 2008.

[20] Disney. BRDF Explorer Tool. [web page] `http://www.disneyanimation.com/technology/brdf.html`, 2013.

[21] Arne Dür. On the Ward model for global illumination. Preprint, September 2004.

[22] Frédéric Drago, Karol Myszkowski, Thomas Annen, and Norishige Chiba. Adaptive logarithmic mapping for displaying high contrast scenes. *Computer Graphics Forum*, 22(3):419–426, 2003.

[23] Mohamed Farouk, Ibrahim El-Rifai, Shady El-Tayar, Hisham El-Shishiny, Mohamed Hosny, Mohamed El-Rayes, Jose Gomes, Frank Giordano, Holly Rushmeier, Fausto Bernardini, and Karen Magerlein. Scanning and processing 3D objects for web display. *3D Digital Imaging and Modeling, International Conference on*, 0:310, 2003.

[24] D.B. Goldman, B. Curless, A. Hertzmann, and S.M. Seitz. Shape and spatially-varying BRDFs from photometric stereo. In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, volume 1, pages 341 – 348 Vol. 1, 2005.

[25] Alexandrino Gonçalves, Luís Magalhães, João Moura, and Alan Chalmers. High dynamic range - a gateway for predictive ancient lighting. *Journal on Computing and Cultural Heritage (JOCCH)*, 2(1):1–20, 2009.

[26] H. Gouraud. Continuous shading of curved surfaces. *IEEE Transactions on Computers*, C-20(6):623–629, 1971.

[27] Tom Haber, Christian Fuchs, Philippe Bekaert, Hans-Peter Seidel, Michael Goesele, and Hendrik Lensch. Relighting objects from image collections. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2009)*, pages 627–634, 2009.

[28] Xiao D. He, Kenneth E. Torrance, François X. Sillion, and Donald P. Greenberg. A comprehensive physical model for light reflection. In *Proceedings of the 18th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '91, pages 175–186, New York, NY, USA, 1991. ACM.

[29] Eugene Hecht. *Optics*. Addison-Wesley, 4th edition, 2002.

[30] Matthias B. Hullin, Johannes Hanika, Boris Ajdin, Hans-Peter Seidel, Jan Kautz, and Hendrik P. A. Lensch. Acquisition and analysis of bispectral bidirectional reflectance and reradiation distribution functions. In *ACM SIGGRAPH 2010 papers*, SIGGRAPH '10, pages 97:1–97:7, New York, NY, USA, 2010. ACM.

[31] Tobias Häußler. Mobile reflectance estimation. Master's thesis, Universität Ulm, Fakultät für Ingenieurwissenschaften und Informatik, 2012.

[32] Imago. 3D Virtual Museum. [web page] `http://www.imago.ufpr.br/Museu/en_index.html`, 2013.

[33] ISO 14524. Photography - electronic still-picture cameras - methods for measuring opto-electronic conversion functions (OECFs), 1999.

[34] Scott Kirkpatrick, Charles Daniel Gelatt, and Mario P. Vecchi. Optimization by simulated annealing. *Science*, 220:671–680, 1983.

[35] Andreas Koschan. A comparative study on color edge detection. In *Asian Conference on Computer Vision*, pages 574–578, 1995.

[36] Ryo Kurazume, Ko Nishino, Zhengyou Zhang, and Katsushi Ikeuchi. Simultaneous 2D images and 3D geometric model registration for texture mapping utilizing reflectance attribute. In *Asian Conference on Computer Vision*, pages 99–106, 2002.

[37] Eric P. F. Lafortune, Sing-Choong Foo, Kenneth E. Torrance, and Donald P. Greenberg. Non-linear approximation of reflectance functions. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '97, pages 117–126, New York, NY, USA, 1997. ACM Press/Addison-Wesley Publishing Co.

[38] Pascal Lagger and Pascal Fua. Retrieving multiple light sources in the presence of specular reflections and texture. *Computer Vision and Image Understanding*, 111:207–218, 2008.

[39] Gregory Ward Larson. Logluv encoding for full-gamut, high-dynamic range images. *Journal of Graphics Tools*, 3:15–31, 1998.

[40] Charles L. Lawson and Richard J. Hanson. *Solving least squares problems*, volume 15 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), 1995. Revised reprint of the 1974 original.

[41] Hendrik P. A. Lensch. *Efficient, Image-Based Appearance Acquisition of Real-World Objects*. PhD thesis, Max-Planck-Institut für Informatik, 2003.

[42] Hendrik P. A. Lensch, Michael Goesele, Jan Kautz, Wolfgang Heidrich, and Hans-Peter Seidel. Image-based reconstruction of spatially varying materials. In *Proceedings of the 12th Eurographics Workshop on Rendering Techniques*, pages 103–114, London, UK, 2001. Springer-Verlag.

[43] Hendrik P. A. Lensch, Jan Kautz, Michael Goesele, Wolfgang Heidrich, and Hans-Peter Seidel. Image-based reconstruction of spatial appearance and geometric detail. *ACM Transactions on Graphics (TOG)*, 22:234–257, 2003.

[44] Hendrik P. A. Lensch, Jochen Lang, Asla M. Sá, and Hans-Peter Seidel. Planned sampling of spatially varying BRDFs. In Pere Brunet and Dieter W. Fellner, editors, *EUROGRAPHICS 2003 (EUROGRAPHICS-03) : the European Association for Computer Graphics, 24th Annual Conference*, pages 473–482, Granada, Spain, 2003. Blackwell.

[45] Marc Levoy, Kari Pulli, Brian Curless, Szymon Rusinkiewicz, David Koller, Lucas Pereira, Matt Ginzton, Sean Anderson, James Davis, Jeremy Ginsberg, Jonathan Shade, and Duane Fulk. The digital Michelangelo project: 3D scanning of large statues. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '00, pages 131–144, New York, NY, USA, 2000. ACM Press/Addison-Wesley Publishing Co.

[46] Shiying Li, Yoshitsugu Manabe, and Kunihiro Chihara. Accurately estimating reflectance parameters for color and gloss reproduction. *Computer Vision and Image Understanding*, 113:308–316, 2009.

[47] M.I.A. Lourakis. levmar: Levenberg-marquardt nonlinear least squares algorithms in C/C++. [web page] `http://www.ics.forth.gr/~lourakis/levmar/`, 2013.

[48] Takashi Machida and Naokazu Yokoya. Virtualizing real objects with surface interreflections. In *Multimedia and Expo, 2002. ICME '02. Proceedings. 2002 IEEE International Conference on*, volume 1, pages 833 – 836 vol.1, 2002.

[49] Takashi Machida, Naokazu Yokoya, and Haruo Takemura. Surface reflectance modeling of real objects with interreflections. In *Proceedings of the Ninth IEEE International Conference on Computer Vision - Volume 2*, ICCV '03, pages 170–177, Washington, DC, USA, 2003. IEEE Computer Society.

[50] Stephen R. Marschner, Stephen H. Westin, Eric P. Lafortune, Kenneth E. Torrance, and Donald P. Greenberg. Image-based BRDF measurement including human skin. In *Rendering Techniques, Proceedings of the Eurographics Workshop*, pages 131–144, 1999.

[51] Wojciech Matusik, Hanspeter Pfister, Matt Brand, and Leonard McMillan. A data-driven reflectance model. *ACM Transactions on Graphics (TOG)*, 22:759–769, 2003.

[52] Caroline M. Mendes, Leonardo Gomes, Beatriz T. Andrade, Wesley A. Castelluber, Ronaldo S. Alburnio, Olga R. P. Bellon, and Luciano Silva. Preservação digital do

acervo nacional: de artes indígenas até obras coloniais de Antônio Francisco Lisboa, o Aleijadinho. In *Special Session on Works in Progress. 23rd Conference on Graphics, Patterns and Images*, 2010.

[53] S.K. Naik and C.A. Murthy. Standardization of edge magnitude in color images. *IEEE Transactions on Image Processing*, 15(9):2588–2595, 2006.

[54] Addy Ngan, Frédo Durand, and Wojciech Matusik. Experimental analysis of BRDF models. In *Proceedings of the Eurographics Symposium on Rendering*, pages 117–226. Eurographics Association, 2005.

[55] Addy Ngan, Frédo Durand, and Wojciech Matusik. Experimental analysis of BRDF models - suplemental material, 2005.

[56] Wai Kit Addy Ngan. *Acquisition and Modeling of Material Appearance*. PhD thesis, Massachusetts Institute of Technology, 2006.

[57] Matt Pharr and Greg Humphreys. *Physically Based Rendering, Second Edition: From Theory To Implementation*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2nd edition, 2010.

[58] Point Grey Research. Ladybug 3 technical reference manual. [web page] `http://www.ptgrey.com/products/ladybug3/`, 2013.

[59] Holly Rushmeier and Fausto Bernardini. Computing consistent normals and colors from photometric data. pages 99–108, 1999.

[60] Joaquim Salvi, Xavier Armangue, and Joan Batlle. A comparative review of camera calibrating methods with accuracy evaluation. *Pattern Recognition*, 35(7):1617–1635, 2002.

[61] Yoichi Sato, Mark D. Wheeler, and Katsushi Ikeuchi. Object shape and reflectance modeling from observation. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '97, pages 379–387, New York, NY, USA, 1997. ACM Press/Addison-Wesley Publishing Co.

[62] Christophe Schlick. An inexpensive BRDF model for physically-based rendering. *Computer Graphics Forum*, 13:233–246, 1994.

[63] Min-Zhi Shao and Norman I. Badler. Spherical sampling by Archimedes' Theorem. Technical Report MS-CIS-96-02, Computer and Information Science Department, University of Pennsylvania, 1996.

[64] Li Shen and H. Takemura. Spatial reflectance recovery under complex illumination from sparse images. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 2, pages 1833 – 1838, 2006.

[65] Noah Snavely, Steven M. Seitz, and Richard Szeliski. Photo tourism: exploring photo collections in 3D. *ACM Transactions on Graphics (TOG)*, 25(3):835–846, July 2006.

[66] Phil Torr and Andrew Zisserman. Robust computation and parametrization of multiple view relations. *International Conference on Computer Vision*, pages 727–732, 1998.

[67] Kenneth E. Torrance and E. M. Sparrow. Theory for off-specular reflection from roughened surfaces. *Journal of The Optical Society of America*, 57, 1967.

[68] TyphoonLabs' OpenGL Shader Designer. [web page] `http://www.opengl.org/sdk/tools/ShaderDesigner/`, 2013.

[69] Unesco. Sanctuary of Bom Jesus do Congonhas. [web page] `http://whc.unesco.org/en/list/334`, 2013.

[70] George Vogiatzis, Paolo Favaro, and Roberto Cipolla. Using frontier points to recover shape, reflectance and illumination. *IEEE International Conference on Computer Vision*, 1:228–235, 2005.

[71] Alexandre Vrubel. Reconstrução digital de objetos com scanners 3D e triangulação laser. Master's thesis, Universidade Federal do Paraná, Campus Centro Politécnico, Curitiba, Paraná, 2008.

[72] Alexandre Vrubel, Olga R. P. Bellon, and Luciano Silva. A 3D reconstruction pipeline for digital preservation. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 2687–2694, 2009.

[73] Greg Ward. *Real pixels, Graphics Gems II.* Academic Press, 1991.

[74] Gregory J. Ward. Measuring and modeling anisotropic reflection. *ACM SIGGRAPH Computer Graphics*, 26:265–272, 1992.

[75] R. Peter Weistroffer, Kristen R. Walcott, Greg Humphreys, and Jason Lawrence. Efficient basis decomposition for scattered reflectance data. In *EGSR07: Proceedings of the Eurographics Symposium on Rendering*, Grenoble, France, June 2007.

[76] Tim Weyrich, Jason Lawrence, Hendrik P. A. Lensch, Szymon Rusinkiewicz, and Todd Zickler. Principles of appearance acquisition and representation. *Foundations and Trends in Computer Graphics and Vision*, 4(2):75–191, February 2009.

[77] Chanki Yu, Yongduek Seo, and Sang Wook Lee. Global optimization for estimating a BRDF with multiple specular lobes. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 319 –326, 2010.